The Role of Linker Histone Globular Domains in Chromatosome Formation

by

Chang-Hui Shen

Ph.D. Thesis

University of Edinburgh

1997

Declaration of Originality

I declare that, unless otherwise stated, this thesis represents my own work and was composed by me.

Chang-Hui Shen

To my parents and Yu-Tsui

Acknowledgements

From the bottom of heart, I would like to express my thanks to my supervisor, Dr. James Allan, for help, constructive advice and leading me to a knowledge of chromatin throughout the past three years.

Thanks are also due to Dr. Colin Davey for his expert assistance and technical advice, and for discussion and criticism of this manuscript.

My appreciation are also due to Dr. Richard Meehan and Dr. Sari Pennings for their invaluable suggestions and advice during my course of study.

I am also grateful to Dr. David Apps and Dr.Sutherland Maciver for providing me with the use of SW41 rotor and Fluorescence spectrometer for the study of GH1 expression.

Finally, I would like to thank Alastair, George, Venkat, Nick, Melanie and all the other members of Biochemistry Department for their friendship and help during my stay in Scotland.

Abstract

Expression of the globular domain of the chicken linker histone H1 (GH1) in *E.coli* has the potential to allow the study of chromatin structure. In a T7 polymerase expression system only very low levels of the wild type chicken GH1 are produced. However, mutating the third base in each of the first 4 codons of GH1, from G or C to A or T, substantially raised expression levels. To further investigate this observation, 15 mutants, comprising all possible combinations of the altered first 4 codons, were constructed and their expression was compared to that of the wild type GH1. Expression levels were found to vary over a 300 fold range. The results indicate that increased GH1 expression (i) is not a function of mRNA abundance as each mutant was transcribed with equal efficiency, (ii) does not depend upon codon usage and (iii) is not correlated to changes in intramolecular mRNA secondary structure *per se* or to its potential influence upon the structure of the Shine-Dalgarno sequence in the translational initiation region of the mRNAs. However, a striking complementarity between the sequence +5 to +19 of wild type GH1 mRNA and a region of the 16S rRNA molecule (1526 to 1510) located only a few nucleotides 5' of the 16S anti-Shine-Dalgarno sequence was noted. This observation could indicate that the potential to form a duplex between the first 4 codons of GH1 mRNA and the 16S rRNA molecule may be the feature responsible for modulating GH1 expression in *E.coli*. The fact that those mutations which increase GH1 expression, substantially decrease the likelihood of hybrid formation support this proposal. Furthermore, the expression of Green Fluorescence Protein (GFP) was greatly inhibited when an oligonucleotide which is homologus to the sequence +4 to +19 of wild type GH1 mRNA was inserted immediately after the start codon of GFP.

The chicken β-globin genes are expressed in a precise temporal and spatial pattern during development. The mechanisms which regulate their expression are not yet fully understood. However, positioned nucleosomes, which contribute to both the structure and the function of the chromatin fiber, could play a decisive role in controlling their expression. Monomer extension, a technique designed to map the precise translational positions adopted by core histone octamers reconstituted onto

long DNA fragments, has been used to study the chicken β-globin gene region and to provide the first long-range nucleosome positioning map for an entire, contiguous gene region (Davey *et al.* (1995), PNAS, 92, 11210-11214). Using the same method, I have determined the translational positions adopted by core histone octamers, reconstituted onto a 1.5 kb stretch of the chicken adult β-globin gene, after titration with the globular domains of linker histone H1 and H5, thus providing a chromatosome map. The results indicate that (i) linker histone globular domains display extensive variation in their binding affinity for different positioned nucleosomes, (ii) the extra 20 bp protected in a chromatosome is most often symmetrically distributed with respect to the core particle structure (10 + 10 bp), although there are exceptions (15 + 5 bp or 20 + 0 bp) to this rule, (iii) The addition of GH1/GH5 can not alter nucleosome positioning sites, (iv) core particle and chromatosome positioning are not influenced by the temperature at which final stages of reconstitution are carried out and (v) in respect of the above properties, there is little difference between the abilities of GH1 and GH5 to form chromatosome. These results suggest that the affinities of GH1 and GH5 for different nucleosomes may depend on the DNA content at the chromatosome boundary.

Contents

Chapter 2 General materials and methods

List of Figures

List of Tables

List of Abbreviations

Å	angstroem 1Å=10^{-10}m
ADB	anti-downstream box
ASD	anti- Shine-Dalgarno
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
DB	downstream box
dATP	2`-deoxyadenosine 5`-triphosphate
dCTP	2`-deoxycytidine 5`-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	2`-deoxyguanosine 5`-triphosphate
DNaseI	deoxyribonuclease I
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
dTTP	2`-deoxythymidine 5`-triphosphate
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
g	gravity acceleration
GH1	globular domain of linker histone H1
GH5	globular domain of linker histone H5
GTE	glucose/Tris/EDTA
GDP	guanosine diphosphate
GFP	green fluorescent protein
GTP	guanosine triphosphate
hr	hour
IPTG	isopropyl-B-D-1-thiogalactopyranosid
Kbp	kilo base
Kd	kilo dalton
Lc	size standard included C sequencing reaction of M13mp18

Lt	size standard included T sequencing reaction of M13mp18
mg	1x10-3 g
min	minutes
MNase	micrococcal nuclease
mRNA	messenger ribonucleic acid
MW	molecular weight [g/mol]
nt	nucleotide
OD	optical density
OD_{600}	optical absorbance at 600 nm.
PEG	polyethylene glycol
pmol	pico mole
PMSF	phenylmethylsulphonyl fluoride
PVP	polyvinylpyrolidone
RBS	ribosome binding site
rpm	rounds per minutes
rRNA	ribosomal ribonucleic acid
SD	Shine-Dalgarno
SDS	sodium dodecyl sulphate
sec	seconds
SSC	sodium chloride/sodium citrate
ssDNA	single stranded DNA
TBE	Tris/Borate/EDTA
TE	Tris/EDTA
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
TIR	translation initiation region
tRNA	transfer ribonucleic acid
UV	ultraviolet light
V	Volts
W	Watts
ΔG^0	minimum free energy
µg	1x10-6 g
µl	1x10-6 l

0GH146	core particle DNA derived from reconstitutes lacking linker histone globular domain
1GH1-146	core particle DNA derived from GH1-containing reconstitutes
1GH1-168	chromatosomal DNA derived from GH1-containing reconstitutes
1GH5-146	core particle DNA derived from GH5-containing reconstitutes
1GH5-168	chromatosomal DNA derived from GH5-containing reconstitutes
1H1-146	core particle DNA derived from H1-containing reconstitutes
1H1-168	chromatosomal DNA derived from H1-containing reconstitutes
1H5-146	core particle DNA derived from H5-containing reconstitutes
1H5-168	chromatosomal DNA derived from H5-containing reconstitutes

Chapter 1 Introduction

1.1 Overview

An organism does not pass on a simulacrum of itself to the next generation, but instead provides it with genetic material containing the information needed to construct a progeny organism. The laws of genetic inheritance were first discovered by Mendel who, from his analysis of pea genetics in 1865, defined a particular factor that could be passed unchanged from parent to progeny. However, it was not until 1957, when Ingram showed direct proof that a gene is actually responsible for controlling the structure of a protein, that people realised the importance of gene functions. In one sense the era of mechanistic studies of gene expression began with this observation.

Although the nucleus constitutes only 5% of the volume of an eukaryotic cell, an immense length of DNA has to be packaged into this small volume. For example, the smallest human chromosome contains 4.6 X 10^7 bp of DNA (≈1.4 cm) which is packed into a 2 μm long chromosome. Therefore, the packing ratio of DNA in the chromosome can be as great as 7000. DNA in a chromosome is maintained in the form of an organised structure – chromatin, and this nucleoprotein complex constitutes the template for the processing of genetic information manifest as transcription, replication and recombination. Chromatin is predominately composed of DNA, histone proteins, and non-histone proteins. In general, the DNA is wrapped around the core histones to form the fundamental chromatin unit – the core particle. Linker histones interact with core particles, and their associated linker DNA, to form a more folded nucleosome and to promote the formation of higher-order levels of chromatin structure from strings of nucleosomes. Experiments carried out over the past decade have shown that changes in chromatin structure between a folded or unfolded state may dominate the control of gene expression. Thus, it is now clear

that the processes by which the eukaryotic genome becomes transcriptionally active require not only transcription factors, but also cooperation with histones and with cofactors that help to disrupt nucleosomes and chromatin structure. Therefore, it is likely that linker histones play a crucial role in the regulation of gene expression by virtue of their ability to facilitate and regulate the organisation of nucleosomes and the degree to which the fibre is folded.

As the structure of the chromatin fibre changes to allow gene activation, the challenge nowadays is to understand the detailed nature of the different chromatin structures involved and to determine how the dynamic transitions between these structures are effected. In terms of their involvement in the regulation of eukaryotic gene expression, it is critical that we discern the functional and architectural role of linker histones so that we can fully understand the developmental and cellular regulation of embryo.

1.2 Chromatin Structure

In the eukaryotic genome, DNA is packaged inside nuclei by association with histone proteins to form a nucleoprotein complex known as chromatin. This complex consists of roughly a 2:1 mass ratio of protein to DNA and a 1:1 mass ratio of histones to DNA. The basic subunit of chromatin is the nucleosome, which comprises two molecules of each of the core histones H2A, H2B, H3 and H4 (the histone octamer) around which is wrapped 1.75 turns DNA. In higher eukaryotic cells, linker histones H1 or H5 bind to the nucleosome core particle to form the chromatosome and promote the organisation of nucleosomes into the 30 nm filament. This packaging of the DNA into chromatin, which occurs at a series of different levels, provides the compaction and organisation of the DNA required to accommodate processes such as transcription, recombination, replication, and mitosis.

1.2.1 Histone proteins

The primary proteins mediating the folding of DNA into chromatin are the histones. Since histones can be removed from DNA by high salt concentrations, the major interactions between DNA and the core histones appear to be electrostatic in nature. All of the core histones are small basic proteins. Most histones have three structural domains, a central structured globular domain and an amino- and carboxyl-terminal flexible basic extension or arm (Table 1-1). All of the core histones contain relatively large amounts of lysine and arginine (van Holde, 1988). The extended histone-fold domain at the carboxyl terminal end of the protein, which is involved in histone-histone and histone-DNA interactions, is predominantly α-helix, with a long central helix bordered on each side by a loop segment and a shorter helix (Pruss *et al.*, 1995). Experiments in which the sequences of the amino terminal tails of core histones were altered indicate that the tails play a part in regulating the transcriptional activation or repression of specific genes (Grunstein *et al.*, 1992). Modifications of these tail domains, including depletion of histones H2A/H2B, proteolytic removal or hyperacetylation of the amino-terminal tails also influence transcription factor access to nucleosomal DNA (Hayes & Wolffe, 1992a; Lee *et al.*, 1993; Chen & Workman, 1994; Vettesse-Dadey *et al.*, 1994; Godde *et al.*, 1995). Therefore, the basic tail domains may serve a regulatory role in modulating the structure of chromatin.

1.2.2 Nucleosome core particle

In the early stages of chromatin structure research, the existence of the nucleosome, the repeat unit of chromatin, was first demonstrated by nuclease digestion (Williamson, 1970; Hewish & Burgoyne, 1973) and electron microscopy of chromatin (Noll, 1974; Oudet *et al.*, 1975; Finch *et al.*, 1975). As the arginine-rich histones H3 and H4 exist in solution as tetramers (Kornberg & Thomas, 1974) and the lysine-rich histones H2A and H2B as dimers (Kornberg & Thomas, 1974), it was proposed (Kornberg, 1974) that the fundamental unit of organisation in chromatin

Core histone	Number of amino acids	Length of amino-terminal tail	Length of carboxy-terminal tail
H4	102	26	2
H3	135	44	3
H2A	128	16	11
H2B	122	32	3

Table 1-1 Number of amino-acid residues for each core histone and the length of their tails. (Data are from Luger *et al.*, 1997)

was a nucleoprotein particle containing an octameric core of histones surrounded by about 200 base-pairs of DNA, the composition of the octamer being $(H3)_2(H4)_2(H2A)_2(H2B)_2$. Subsequent experiments employing chemical cross-linking (Thomas & Kornberg, 1975) and hydrodynamic analyses (Noll & Kornberg, 1977) provided support for an octameric core structure, which was also consistent with the fact that core histones are present in equimolar amounts in chromatin. In 1975, Oudet *et al.* proposed that the chromatin fibre consists of a contiguous array of such particles, which appeared like beads on a string in the electron microscope, and proposed the term nucleosome (Oudet *et al.*, 1975). Nuclease digestion studies have shown that extensive digestion of the nucleosome produces a particle comprising a set of eight histone molecules and 146 bp of DNA. This basic structure is termed the core particle and seems to be invariable in all cell types (Mirzabekov *et al.*, 1978; Lutter, 1979).

1.2.3 Structure of nucleosome core particle

The structure of the nucleosome core particle was first described at 20 Å resolution from X-ray diffraction and electron microscopy studies (Finch *et al.*, 1977). This work suggested the core to be a flat, somewhat wedge shaped particle of approximate dimensions 110 X 110 X 57 Å, markedly divided into two layers (bipartite) and with the DNA wound into about $1^3/_4$ turns of a flat superhelix of a pitch about 28 Å. Klug *et al.* (1980) used electron microscopy and image reconstruction to produce a 22 Å three dimensional density map which suggested that the histone core has a roughly circular outline of approximate diameter 70 Å and length 55 Å, a two-fold axis of symmetry and the overall shape of a left-handed helical spool on which to wind about two turns of a superhelix of DNA. They proposed that the $(H3)_2(H4)_2$ tetramer forms a dislocated disk which defines the surface for the central turn of DNA, while the two H2A-H2B dimers lie one on each face, each associated with about half a turn of additional DNA. The division of the octamer into two H2A-H2B dimers residing on opposite faces of an H3-H4 tetramer

agrees well with the physicochemical solution studies which describe the core histone octamer as a system of three thermodynamic entities in freely reversible chemical equilibrium (Eickbush & Moudrianakis, 1978). The model they proposed is the foundation of nucleosomal core histone octamer structure and was called the MRC model (Arents *et al.*, 1991).

The core particle model was further refined from a 7 Å resolution crystal structure of the nucleosome core particle (Richmond *et al.*, 1984). The low-resolution structure of the core particle revealed that the histone octamer forms a helical ramp around which is wrapped 1.7 turns of a left-handed DNA superhelix. This analysis also demonstrated many structural features associated with the core particle: (i) the DNA is not bent uniformly into the superhelix, but exhibits several regions of tight bending or possible kinking, adjacent to points of strong contact with histones H3 and H4; (ii) the histone-DNA interactions occur on the inside of the superhelix; (iii) the central turn of the DNA superhelix and the H3-H4 tetramer have dyad symmetry, but the H2A-H2B dimers show departures from symmetry due to interparticle associations. The use of multiple heavy-atom compounds in the multiple isomorphous replacement method for the phase determination process was an important step in the structure solution, avoiding the use of models in the phasing of the X-ray data (Richmond *et al.*, 1984; O'Halloran *et al.*, 1987). Although crystals were treated with 1,6-hexanediol to partially dehydrate the crystals, a fully hydrated 9 Å structure of the core particle in the absence of alcohol showed that the basic structural features, including the sharp turns in the DNA at positions ±1 and ±4 turns from the DNA centre (dyad) were retained (Struck *et al.*, 1992).

Our understanding of the structure of the octameric histone core was greatly improved by X-ray crystallography of the octamer to a resolution of 3.1 Å in the absence of DNA (Arents *et al.*, 1991). As shown in Figure 1-1A, the histone octamer is a cylindrical wedge with a persistently curving outer surface that resembles a left-handed helical ramp. The outer dimensions of the histone octamer are consistent with the MRC model (Finch *et al.*, 1977; Klug *et al.*, 1980). In addition, none of the

four core histones is compacted into a single globular domain. As shown in Figure 1-1B, each histone chain is folded in a rather elongated fashion; upon assembly into their physiological subunits, the domains of the folded polypeptides interdigitate extensively rather than each chain occupying a unique and contiguous segment on the surface of the octamer. This arrangement generates the potential for several noncontiguous contacts between each of the four polypeptides and the DNA helix as it winds its path around the octamer (Figure 1-1A). Consequently, the resulting superhelical surface has a complex topography to which each histone chain contributes a minimum of two distinct, separate domains. Starting at the outermost point along the histone cylinder and travelling in a spiral path of 28 Å pitch toward the two-fold axis at the front of the tetramer, the structured portions of the histones emerge on the surface in the following order: $H2A^1$, $H3^2$, $H2A^1$, $H2B^1$, $H2A^1$, $H2B^1$, $H4^1$, $H3^1$, $H4^1$, and finally $H3^1$-$H3^2$ overlapping at the zero position (Shick *et al.*, 1980; Bavykin *et al.*,1990; Arents & Moudrianakis, 1993).

The contact between DNA and the surface of the octamer at the point where the two molecules of H3 meet at the molecular two-fold axis is referred to as zero (Klug *et al.*, 1980; van Holde, 1988). The turns of the advancing DNA double helix are numbered, from one to the other end of the core particle, from -7 to +7, passing through the zero point. In total this 14-turn DNA helix makes contact with 14 patches on the octamer surface (Arents & Moudrianakis, 1993). For the central 12 patches (from -5.5 to +5.5) the contact surfaces are rather large, while at the outer 2 patches (at ±6.5) the contact areas are significantly smaller.

The extrafold N-terminal helices of H3 and H2A and the C-terminal helix of H2B appear to make significant contributions to DNA binding. There are three types of amino acids at the DNA binding sites: (i) positively charged residues - lysines or arginines; from Figure 1-1A, it is clearly seen that lysines and arginines interact with phosphate groups along the entire DNA binding surface; (ii) hydroxyl-containing residues - serines or threonines; for example the first six imaged residues of H2A (residues 15-20) and the last seven residues of H2B (residues119-125) contain

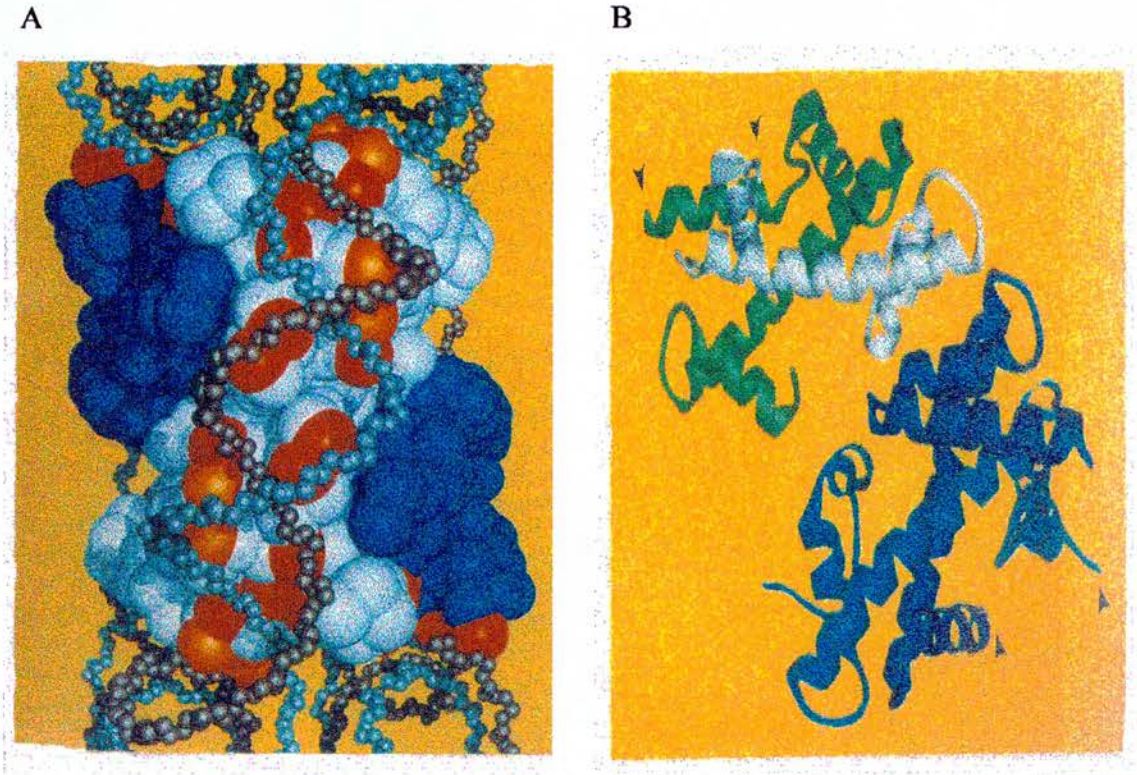


Figure 1-1 A. Histone octamer footprints on the cylindrical surface and complementarity between positive charges on the core protein and the DNA phosphodiester backbones. The (H3-H4) tetramer is white and the (H2A-H2B) dimers are blue. The C-α atoms of lysines and arginines on the cylidrical surface are indicated in red. B. Ribbon model of the four histones. H4 is white, H3 is green, H2A is light blue and H2B is dark blue. The amino end of each chain is marked by an arrow. (Taken from Moudrianakis & Arents, 1993).

exclusively positively charged or hydroxyl-containing amino acids and are likely to be important in DNA binding; (iii) hydrophobic residues – there are only 7 residues found in the 4 core histones : Ile-79 (H2A), Ile-39 (H2B), Tyr-40 (H2B), Tyr-42 (H2B), Leu-65 (H3), Met-120 (H3) and Leu-49 (H4). The primary histone-DNA contacts are balanced on the two sides of the double helix. Each histone chain in a dimer contributes to all three of the DNA binding sites within that dimer, and thus both dimer partners make analogous contributions to the binding of each DNA strand. An individual strand of DNA will interact primarily with one of the histones in two places and with the other histone only once. This pattern suggests that modifications in a single histone might generate changes in the nucleosome at several widely spaced locations (Arents & Moudrianakis, 1993).

Recently, the X-ray crystal structure of the nucleosome core particle has been further resolved in atomic details (~2.8Å) as shown in Figure 1-2A (Luger *et al.*, 1997). This higher resolution structure revealed in detail the assembly of histone proteins and the arrangement of the nucleosomal DNA. It showed that the central histone-fold domains of all core histone proteins have a similar structural motif, α1 helix - L1 loop - α2 helix - L2 loop - α3 helix (Figure 1-2B). The C-terminal halves of the α2 helices and α3 helices are the main determinants of the tetramer and octamer by virtue of their ability to form three, four-helix bundles (Figure 1-2C); more specifically, the α3 helices of H4, H3 and H2B are the principal interaction sites for assembly of tetramer and octamer, and the H2A α3 helix may contribute to the capacity of nucleosome to form higher order structure. These authors also defined two types of DNA-binding sites: (i) α1α1 sites, using both α1 helices to bind the two DNA backbone segments at the centre of the bound DNA stretch; (ii) L1L2 sites, forming L1 and L2 loops and termini of the α2 helices at each end of the histone-pair crescent. The α1α1 DNA-binding sites of the H3-H4 pairs cause an outward bulge in the DNA. Similarly, the adjacent L1L2 and α1α1 DNA-binding sites of the H2A-H2B dimers buckle the DNA outwards. Therefore, the path of DNA superhelix is significantly distorted (sites ±1.5 and ±4 to 5).

A

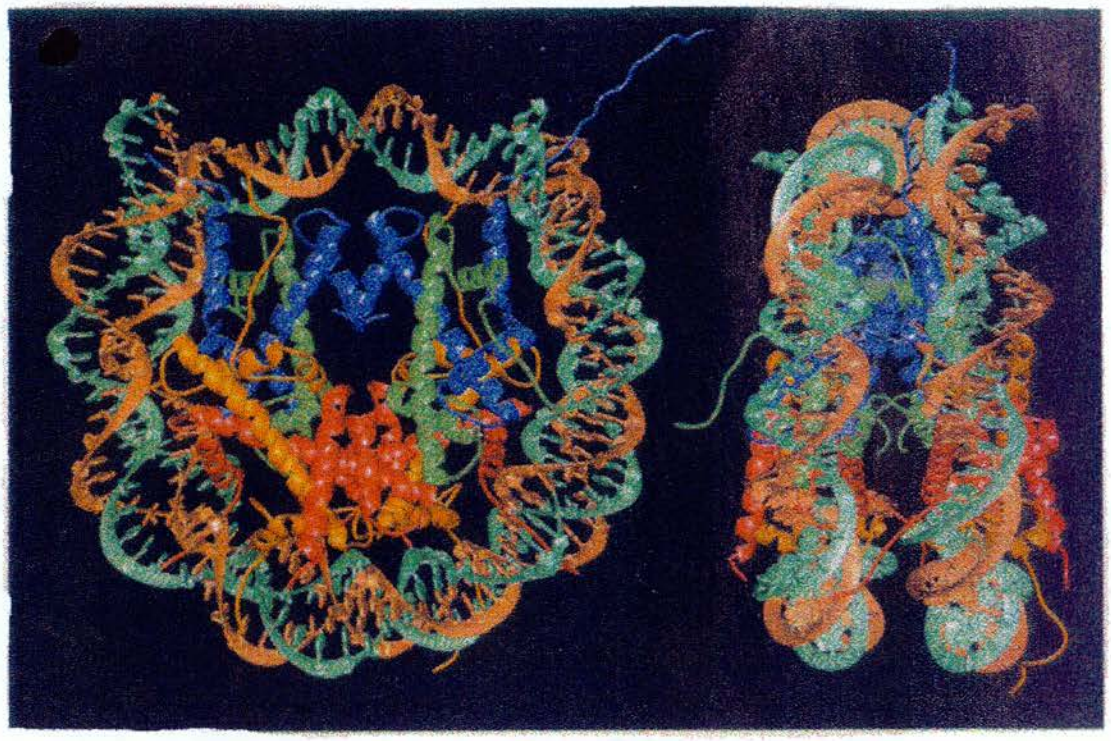
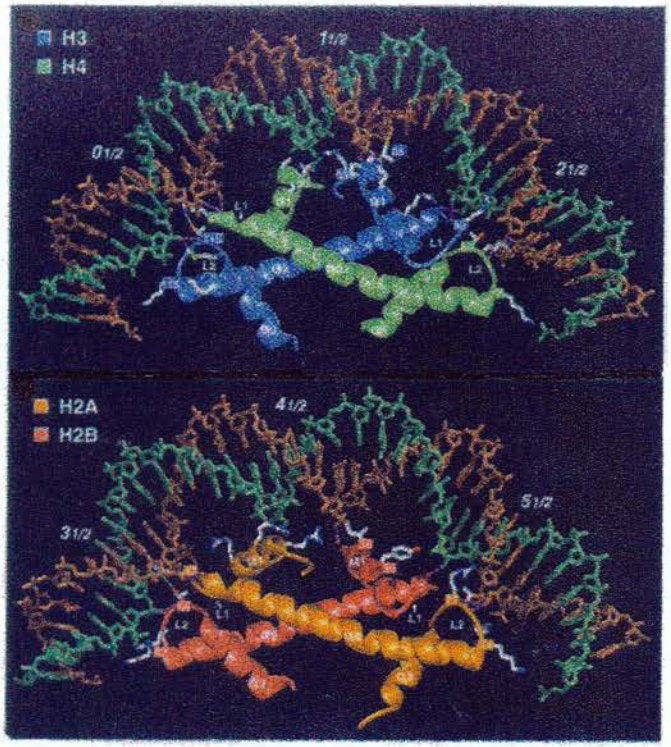


Figure 1-2 A. Nucleosome core particle: left particle is a down view of the DNA superhelix axis and right particle is a perpendicular view. H4 is green, H3 is blue, H2A is yellow, and H2B is red. (Taken from Luger *et al.*, 1997).

B



C

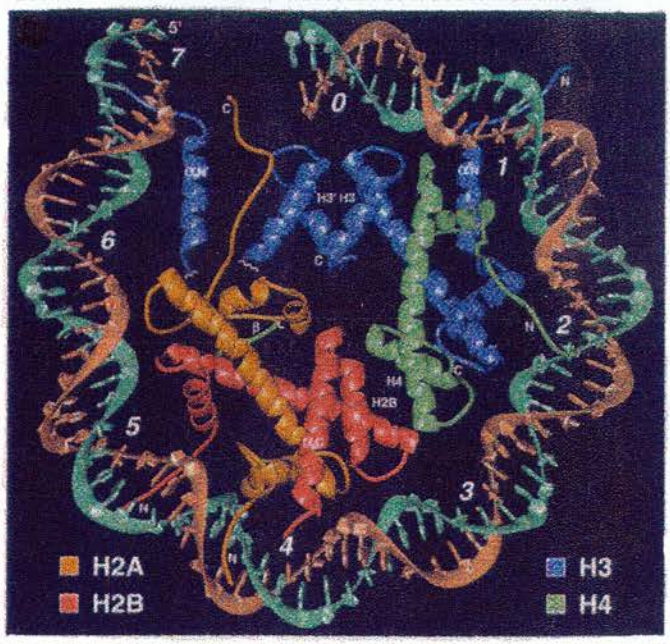


Figure 1-2 B. H3-H4 and H2A-H2B histone-fold pairs. Numbers are superhelix location. C. DNA superhelix axis and nucleosome particle showing DNA distortion and four-helix bundles. (Taken from Luger *et al.*, 1997)

The location and extension of the core histone tails within the core particle are also identified (Figure 1-2A & C). The N-terminal tails of both H3 and H2B pass through the gyres of the DNA superhelix creating a 20-bp periodicity of tails interaction with the minor groove channels. The H2A N-terminal tail is bound to the minor groove on the outside of the superhelix. The basic side chain of the H4 N-terminal tail makes multiple hydrogen bonds and salt bridges with acidic side chain residuals of H2A and H2B.

Hydroxyl radical cleavage of nucleosome cores has revealed that the structure of DNA is different when it is wrapped around the histone octamer compared to when it is free in solution (Hayes *et al.*, 1990; Hayes *et al.*, 1991; Hayes & Wolffe, 1992a). Importantly, on a longer DNA fragment associated with the histone octamer, histone-DNA contacts extend over at least 160 bp, suggesting that two full turns of DNA may be wrapped around the core histones (Hayes *et al.*, 1990; Pruss & Wolffe, 1993). Therefore, the parameters of the double helix wrapped around the core particle could have substantial variation over the length of the superhelix. For example, within the nucleosome core, three helical turns of DNA at the centre of nucleosome have a period of 10.7 bp/turn, whereas outside this region the helical period is 10.0 bp/turn. This DNA structure, which on average is about 10.2 bp/turn, is ideally suited to match the symmetry of repeating DNA binding sites on the surface of the octamer - the number of base-pairs between one DNA binding site and the next being 10.0-11.0 (Arents & Moudrianakis, 1993). The overall helical period of 10.2 bp/turn could create an alignment of minor grooves between superhelix gyres for the passage of the H3 and H2B tails (Luger *et al.*, 1997).

1.2.4 Post-translational modification of core histones

In the core histones, most of the highly conserved sites for reversible post-translational modifications are located in the N termini. Core histones undergo three major post-translational modifications: acetylation, phosphorylation and

ubiquitination. These cell-cycle dependent processes modulate histone-DNA interactions in eukaryotic chromosomes and can thereby affect its structural transitions and spatial rearrangements.

1.2.4.1 Acetylation

Acetylation of the core histones occurs in all animal and plant species and has been correlated with all aspects of DNA processing in eukaryotes: replication, transcription and spermiogenesis (Csordas, 1990). The acetylation reaction involves the transfer of an acetyl group from acetyl coenzyme A onto the ε-amino group of specific lysine residues present in the amino tail of each of the core histones resulting in the neutralisation of a single positive charge (Hong *et al.*, 1993). The pattern of specific lysine residues in the histone tails that are acetylated varies between different species. The sites of modification are the lysine residues of the positively charged amino terminal tails. Acetylation is an energy-intensive, dynamic phenomenon, the steady-state balance of which is mediated by the opposing activities of histone acetyltransferase (HAT) and deacetylase enzyme systems. The HATs are classified operationally as HAT A, which possess a nuclear location and are involved in post-synthetic acetylation of nucleosome core histones, and a connection to transcription. A second class, HAT B, has a cytoplasmic localisation, acetylates free histones rather than nucleosomal histones, and is associated with histone deposition onto replicated DNA. The major HAT systems exhibit different substrate specificities: HAT A's acetylate all four core histones whereas HAT B's primarily acetylates H3 and H4, at lysine positions that are generally different from those modified by HAT A (Turner & O'Neil, 1995; Brownell & Allis, 1996).

Although a 15- to 30-fold enrichment in active sequences has been found in hyperacetylated area of chromatin (Hebbes *et al.*, 1988), the influence of acetylation of the histone tails on chromatin structure is not well defined. It appears that the most significant consequences are for protein-protein interactions, either between nucleosomes, with histone H1 or with non-histone proteins (Stefanovsky *et al.*, 1989; Ausio *et al.*, 1989). It has been suggested that acetylation changes the angle of the

DNA entering/leaving the nucleosome, thereby loosening the binding of H1 and possibly facilitating the displacement of an H2A-H2B dimer that is associated with the ends of the core particle DNA (Thorne *et al.*, 1990). Reconstitution experiments also suggests that core histone acetylation alters the capacity of the H1 histones to condense transcriptionally active/poised chromatin (Ridsdale *et al.*, 1990). Furthermore, the acetylation of H4 allows remodelling of the nucleosome by the incorporation of variants of histones H2A and H2B (Li *et al.*, 1993; Perry *et al.*, 1993). These changes in nucleosome structure could make available binding sites for *trans*-acting factors (Turner, 1993). Acetylation also disrupts specific interactions between the histone tails and non-histone regulators as shown for the yeast silencer and repressor proteins Sir3 and Sir4 (Thompson *et al.*, 1994; Hecht *et al.*, 1995) and Tup-1 (Edmondson *et al.*, 1996) respectively. Thus, histone acetylation may serve as a major means by which chromatin structure may be modulated to accommodate processes such as transcription.

1.2.4.2 Phosphorylation

Although all of the animal histones are capable of being phosphorylated, the preference in transcriptionally active chromatin is for phosphorylation of histones H2A and H3 (van Holde, 1988). As the phosphorylation sites tend to be concentrated in the histone tails, phosphorylation may affect the higher levels of chromatin organisation. Histone H3 is rapidly phosphorylated on serine residues within its basic amino terminal domain. This domain may interact with the ends of DNA in the nucleosomal core particle. Phosphorylation of H3 N-terminal could decrease the $(H3\text{-}H4)_2$ tetramer association constant. Phosphorylating two sites in the H2B globular region reduced (H2A-H2B) dimerization ability. Therefore, phosphorylation may induce loosening of the nucleosome particle and induce a change in nucleosome conformation or higher-order structure. *In vivo*, phosphorylation of H4 and H2A occurs in the cytoplasm shortly after histone synthesis. The phosphorylation of these histones, together with the diacetylation of histone H4, may selectively target them to the molecular chaperones involved in nucleosome assembly at the replication fork (Kaufman & Botchan, 1994).

1.2.4.3 Ubiquitin

Ubiquitin, the most conserved of all eukaryotic proteins, is a very stable, globular 76 amino acid protein. It is covalently attached to H2A and H2B by an isopeptide bond between the C-terminus of ubiquitin and the ε amino group of the target lysine side chain. Thus, ubiquitin forms unusual bifurcated nuclear proteins with H2A and H2B. The functions of these nuclear histone ubiquitinations are not known, however it is believed that ubiquitin may label active or potentially active genes of a particular gene family to prevent the close packaging of nucleosomes in metaphase chromosomes (Mueller *et al.*, 1985b; Nickel *et al.*, 1989).

1.2.5 Structural implication in evolution

The general structure of the histones is similar in both animals and plants. Histones H3 and H4 are two of the most highly conserved proteins. Generally, only 1 or 2 conservative substitutions separate the variants (Brandt & von Holt, 1986), as in the H4s from calf and pea, which differ in sequence by only 2 out of 102 residues (DeLange *et al.*, 1969). The greater range of H2A/H2B variants, particularly variants of H2B (Zweidler, 1984), is typical of plants. The plant H2 variants have higher molecular weights than their animal counterparts and differ in their tail sequences (Spiker, 1982; Klimyuk & Karpenchuk, 1988). The core histone protein sequences have been analysed by phylogenetic trees using the neighbour-joining method. It shows that the core histones have similar reconstructed phylogenies and appear to have co-evolved. H3 and H4 are more conserved than H2A and H2B probably because they have a critical role in nucleosome formation (Thatcher & Gorovsky, 1994). The highly conserved feature of each core histone may be due to the function of each core histone to make numerous interactions, including interactions with its dimer partner, with other components of the octamer, and with DNA, and is therefore subject to a variety of selective pressure (Ramakrishnan, 1995).

By using the quantitative comparisons of the degree of structural similarity between the histone folds of the four core histones, it showed that the chromatin modulations could be facilitated by structure diversification at the level of the histone dimers involved in generating the protein superhelix of the octamer. Also, the N-terminal and C-terminal halves of the histone fold have great similarity (Moudrianakis & Arents, 1993). Therefore, evolution allowed considerable variation in primary structure to accommodate requirements of the higher order structure and interaction with transcription or remodelling factors, but only to the extent that the pattern of the histone fold was preserved (Arents & Moudrianakis, 1995; Luger *et al.*, 1997).

Although the four core histones have little sequence similarity with respect to one another, they may have a common origin. From phylogenetic trees analysis, Thatcher & Gorovsky (1994) suggest that histones which form dimers, H2A-H2B and H3-H4, have very similar trees and appear to have coevolved. H3 and H4 are 10-fold less divergent than H2A and H2B. The low rate of divergence of H3 and H4 may be because the $(H3\text{-}H4)_2$ tetramer is critically involved in DNA binding and nucleosome formation. H2A-H2B dimers, which may interact with non-histone chromatin proteins, transcription factors, and the transcription apparatus, are probably not only permitted but required to evolve to optimise interactions with other proteins as they have evolved along different eukaryotic lineages.

1.2.6 The chromatosome

The next level of chromatin packing above the core particle is the chromatosome which contains about 168 bp of DNA and a molecule of linker histone (Simpson 1978). The chromatosome is a particle which is transiently produced during nuclease digestion of chromatin and its appearance is diagnostic for the presence of linker histone.

1.2.7 Linker histones

Linker histones constitute a family of structurally related proteins which act to stabilise the nucleosome, have an essential role in organising nucleosomes into higher order structures, and may have a role as repressors of transcription (van Holde, 1989).

Typically, linker histones have 3 structural domains : a short, basic unstructured N- terminal domain (20-40 amino acids), a long C-terminal domain (80-140 amino acids) and a folded central globular domain of ~80 amino acids (Hartman *et al.*, 1977; Aviles *et al.*, 1978). The central globular domain is the most conserved region of this molecule compared to the tails.

Reconstitution experiments using either intact H1 or isolated fragments thereof have suggested that the globular domain of linker histone is necessary and sufficient for the protection of the 168 bp chromatosomal DNA (Allan *et al.*, 1980). The globular domains of linker histone are also responsible for positioning the basic N- and C-terminal tails of the H1 molecule so that they can effectively induce higher order structure (Allan *et al.* 1986). However, the ways in which the tails effect the tight packaging of the nucleosome and linker histone to self-associate via their globular domains are still not known.

1.2.7.1 Linker histone variants

The most commonly class of linker histone is H1. There are many H1 variants found within any one species or cell type and they differ in molecular mass, amino acid composition, sequence, and physico-chemical and immunochemical properties (Cole, 1984). It has often been speculated that the pattern of linker histones within a cell is somehow related to the functional properties of that cell. For example, during embryo development in sea urchins the linker histone subtype H1cs, H1α, H1β, and H1γ are observed in differing amounts and ratios as a function of

development stage. Similarly, somatic mouse cells often display different pattern of linker histone content based on the pattern of H1a, H1b, H1c, H1d and H1e.

One of the most intensively studied linker histones is histone H5, which accumulates during the process of terminal differentiation in nucleated erythrocytes. Histone H5 is believed to be specifically involved in chromatin compaction and transcriptional inactivation. A related linker histone, Histone H1°, is found in mammalian tissues showing little cell division. Histone H1° has, therefore, been implicated in the establishment and maintenance of the terminally differentiated state (van Holde, 1989; Zlatanova & Doenecke, 1994). Finally, histone H1t is a tissue-specific variant which only appears in a stage-specific manner during the development of the spermatocyte in the mammalian testis.

1.2.7.2 The structure of the linker histone globular domain

Clore *et al.* (1987) were the first to use distance geometry and restrained molecular dynamics calculations to determine the NMR structure of the globular domain of linker histone H5 (GH5). They found a structural core made up of residues 3-18, 23-34, 37-60, and 71-79 (numbered with respect to the N-terminus of the trypsin-produced globular domain), and two loops composing of residues 19-22 and 61-70. The structure of this core is dominated by three helices (helix I, II, and III). The structure appears to be stabilised by hydrophobic interactions involving Tyr-32, Tyr37 and Phe72 as well as long-chain aliphatic residues, and was observed to have a similar fold to the C-terminal DNA-binding domain of the cAMP receptor protein. Because of the number of lysine residues protected from selective radio-labelling during reductive methylation of histone H5 bound to the nucleosome (Thomas & Wilson, 1986), Clore *et al.* (1987) predicted that the polar residues at the N-terminal end of helix III and helix II may be involved in protein-DNA contacts. Overall, the structure determined by Clore *et al.* (1987) was in agreement with a model in which GH5 sits within a cage of three double helical DNA strands at the entry and exit points of the chromatosome. In this model, the globular domain was proposed to interact with the central DNA strand and the inner surface of the two outer strands.

Recently, the structure of GH5 has been solved to 2.5 Å resolution by multiwavelength anomalous diffraction of crystals of the selenomethionylated protein (Figure 1-3; Ramakrishnan *et al.* 1993). Both the NMR and crystal structures display the same three-helix topology and the assignment of the residues that make up the helices are very similar. The two structures, however, are significantly different in some aspects. There are two molecules, A and B, in the asymmetric crystal unit and each molecule consists of a three-helix bundle (helices I-III), with a B-hairpin at the C-termius. The interface between the A and B monomers consists of helices I and II, and this is a solvent-mediated interaction involving Tyr53, Tyr58 and His57 from both molecules. This may explain the tendency of GH5 to dimerize in solution and suggest that these residues may be responsible for the stabilisation of higher order chromatin structure.

Crane-Robinson & Ptitsyn (1989) compared 24 amino acid sequences for the central globular domain of various linker histones and concluded that one face of the folded globular domain, which contains a cluster of basic residues on the outward facing sides of helices B and C, is the principal DNA binding site. Two opposing faces, orthogonal to the principal site and also containing conserved basic residues, were suggested to be subsidiary DNA binding sites. However, Ramakrishnan *et al.* (1993) noted that, on the basis of the homology between GH5 and CAP (catabolite gene activator protein), the highly conserved residues Lys69 and Arg73 should be positioned to interact with one strand of DNA, and that Lys85 is positioned to interact with the other strand and that these are the primary binding sites (Figure 1-3). Lys40, Arg42, Lys52, and Arg94 make up a putative secondary binding site, and His25 and His62 are likely to be other subsidiary DNA binding sites. By using site-directed mutagenesis, Buckle *et al.* (1992) proved a critical role for Lys85 in determining the effective interaction between GH5 and the nucleosome. Lys52 and Lys69 were also thought to play a major part in the interaction as these residues are protected from chemical modification (Thomas & Wilson, 1986). Recently, Goytisolo *et al.* (1996) have confirmed that at least two of the DNA binding sites on

the GH5 molecule are required to position the globular domain correctly on the nucleosome and to provide chromatosomal stabilisation.

1.2.8 Location of the linker histone globular domain in the chromatosome

On the basis of nuclease digestion experiments, the globular domain of linker histone alone was found to protect linker DNA immediately adjacent to the core structure. Therefore, it is a first degree element for recognition of the nucleosome. It was suggested that the globular domain binds to the nucleosome at the point where the DNA enters and exits the nucleosome, and that the location of GH1/GH5 is orientated directed away from the nucleosomal surface at the dyad axis (Allan *et al.* 1980). As shown in Figure 1-4A, the globular domain of linker histone binds to the outside of the superhelix of DNA at the dyad of the nucleosome and makes three, symmetrical interactions with the nucleosomal DNA which protects 10 bp of linker DNA on either side of the nucleosome core. Many experiments have demonstrated support for this proposal. For example, DNase I footprinting of linker histone-containing dinucleosome complexes showed that the globular domain of linker histones binds directly to the dyad axis of nucleosome (Allan *et al.* 1980; Staynov & Crane-Robinson, 1988).

Recently, an alternative hypothesis based on studies of nucleosomes incorporating specific DNA sequences suggests that a single molecule of linker histone interacts asymmetrically with DNA in the nucleosome (Hayes & Wolffe 1993). From both nuclease protection and histone-DNA crosslinking methodologies, Hayes *et al.* (1994) provided additional evidence for an asymmetric interaction of linker histones with *Xenopus borealis* 5S DNA in a positioned nucleosome. Their results indicated an asymmetry of protection of DNA in the 5S nucleosome with 5 bp and 15 bp being protected at either side of the nucleosome core by the globular domain (Figure 1-4B). By measuring crosslinking of linker histone to nucleosomal DNA, Pruss *et al.* (1996) proposed that GH5 binds to the major groove of the DNA

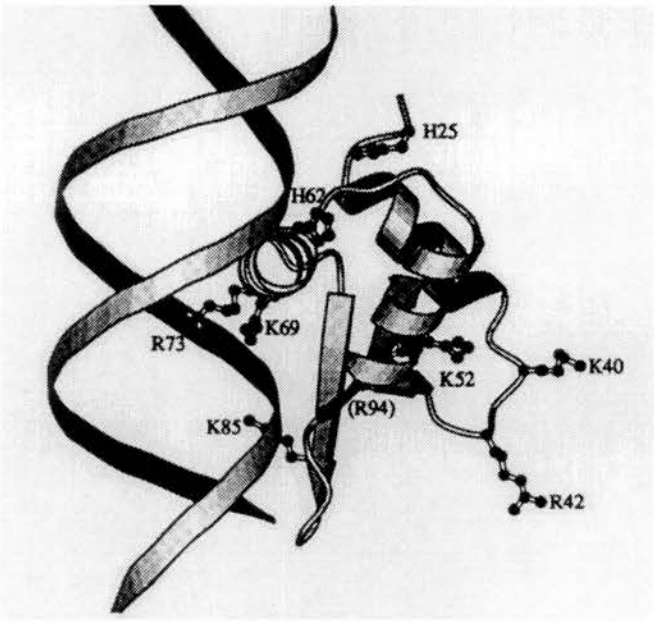


Figure 1-3 Structure of GH5 showing the positions of the basic residues at both sides which contact DNA. (Taken from Goytisolo *et al.*, 1996)

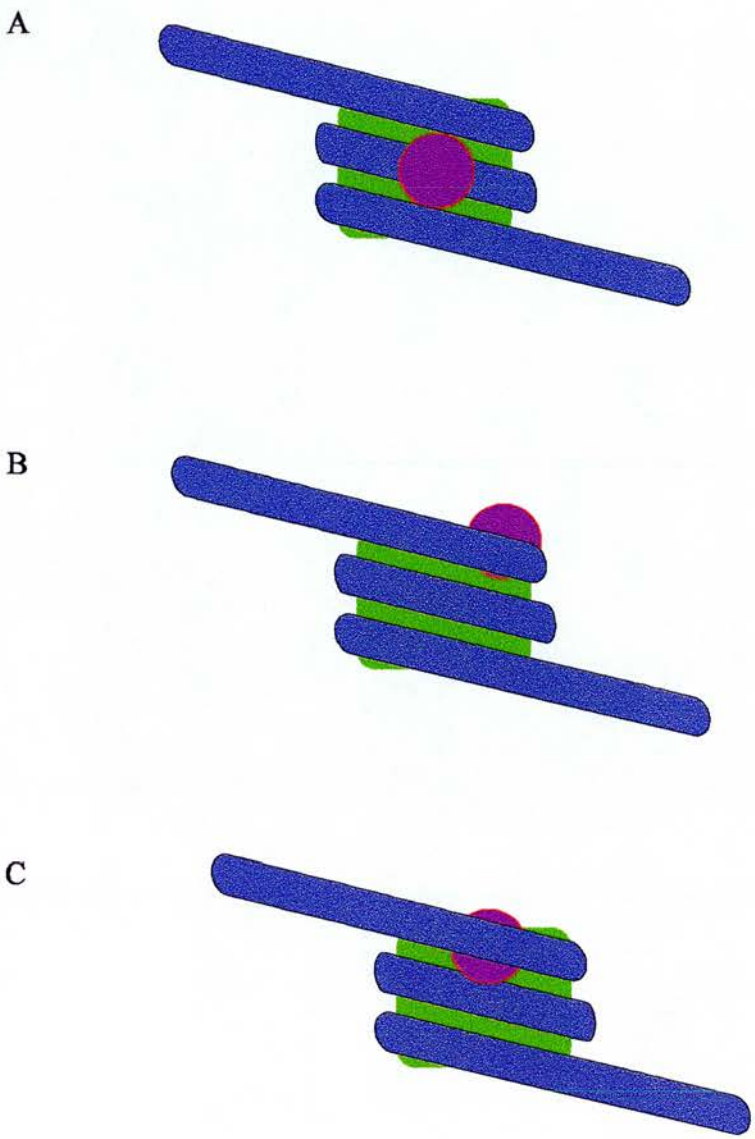


Figure 1-4 Models for the location of the linker histone globular domain on the nucleosome. A. symmetrical model from Allan *et al.* (1980). B. Asymmetrical model from Hayes *et al.* (1993) & Pruss *et al.* (1996). C. Boundary asymmetrical model from Travers & Muyldermans (1996). (Double helix DNA is blue, core histone octamer is green, and linker histone globular domain is purple).

at an open site ~68 bp from the dyad. At this site, GH5 faces in towards the core histones and is in contact with histones H2A and H3. More recently, a site-directed chemical mapping method has been employed which suggested that the location of the linker histone-binding site is located on the inside of the superhelical gyre of DNA, near the periphery of the nucleosome complex, which means that a single globular domain interacts asymmetrically with DNA in the nucleosome (Hayes, 1996). Recently, Eisfeld *et al.* (1997) found that the binding of NF1 to linker DNA in the MMTV promoter was unaffected by incorporation of histone H1 into the MMTV nucleosome particle, thus supporting this proposal.

In addition to the above model, Travers & Muyldermans (1996) have suggested a third possible hypothesis. Their model is based on studies of the sequences of chromatosome DNAs isolated from chicken erythrocytes. They noted that the elements AGGA and AAA/TTT were frequently located at one of the chromatosomal DNA termini and midpoint, respectively. On the assumption that this specificity of sequence location was brought about by linker histones they concluded that the linker histone globular domain contacts chromatosomal DNA termini (Figure 1-4C).

Thus much uncertainty remains as to the location and binding of linker histone globular domain within the chromatosome.

1.2.9 The function and location of linker histones tails

The highly charged tails of linker histones are important for chromatin condensation (Thoma *et al.*, 1983; Allan *et al.*, 1986). The tails are the least conserved parts of the molecule and are thought to be largely unstructured and highly extended in solution, probably because of the electrostatic repulsion between the closely-spaced, positively-charged residues (Lewis & Reams, 1983). However, trypsin digestion of sea urchin sperm H1 in chromatin has suggested that a proline-

free segment of the C-terminal tail immediately adjacent to the globular domain is in an α-helix conformation (Hill *et al.*, 1989) and this may be important for stabilisation of the tail in linker DNA. The tails contain sites employed for cell cycle dependent phosphorylation, a process which is correlated with chromatin condensation.

1.2.9.1 A DNA-binding motif in linker histone - S/TPKK

The role of the H1 tails is thought to be to bind to linker DNA and stabilise chromatin structure (Allan *et al.*, 1986). However, histone H1 is not generally considered to bind to DNA with sequence specificity. Indeed, these molecules are assumed to bind readily to many different sequences of DNA. Nevertheless, H1 histones do exhibit sequence preference arising from the sequence-dependent structural properties of DNA (Churchill & Travers, 1991).

Based on the analysis of the structure of the N-terminal tail of sea urchin sperm histone H1, Suzuki (1989) first identified that repeats of Ser-Pro-Lys(Arg)-Lys(Arg) (SPKK) residues were responsible for the binding of the tail to the minor groove of linker DNA. The SPKK motif was thought to be able to undergo a dynamic conversion between a β-turn structure and a ρ-turn structure (Suzuki, 1988; Suzuki *et al.*, 1993). It was suggested that the β-turn structure could be stabilised by hydrogen bonding between the hydroxyl group and carboxyl group of Ser and the amino group of Lys. These peptide repeats are also found in the tails of H1. For example, the octapeptide KSPKKAKK (8 mer) is present in many histone H1 subtypes, and the imperfect repeat ATPKKSTKKTPKKAKK (16 mer TPKK) occurs in the C-terminus of rat histone H1d (Khadake & Rao, 1997). These studies also suggested that the β-turn structure of the S/TPKK motif is important for binding to the DNA and that the pairing of S/TPKK motifs is a required feature for bringing about DNA and chromatin condensation.

The binding of Hoechst 33258 to DNA can be measured by fluorescence, and the mode of binding has been determined at atomic resolution. The binding of Hoechst 33258 to DNA is inhibited by SPKK repeat motifs, suggesting that the mode

of binding of the SPKK motifs to AT rich DNA may well resemble the minor groove binding characterised for Hoechst 33258 (Churchill & Suzukui, 1989). Histone H1 binds preferentially to SAR (scaffold associated regions)-containing DNA fragments *in vitro* (Kas *et al.*, 1989; Izaurralde *et al.*, 1989). Khadake & Rao (1997) have shown that the motif ATPKKSTKKTPKKAKK (16 mer TPKK), which occurs in the C-terminus of rat histone H1d, can condense SAR DNA fragments effectively. Furthermore, they proposed that within a polynucleosomal fibre, the SAR sequences present in the internucleosomal linker DNA provide high affinity binding sites for histone H1 which can bring about compaction presumably through utilising the SPKK repeat motifs present in the C-terminus of the histone H1. However, it has also been shown that histone H1e from murine erythroleukemic cells and an H1e peptide (25 mer) containing the highly conserved sequence S/TPKKAKKP, binds preferentially to GC-rich regions of DNA (Wellman *et al.*, 1994). In the light of this contradictory evidence, the binding preference of S/TPKK motifs to DNA still remains to be clarified.

The sequence motif SPKK is generally found in the C-terminus of H1, while the sequence motif SPXX, where X can be Ser, Thr, or Ala but rarely a basic residue, is found in gene regulatory proteins (Suzuki, 1989). Examples include Gal 4 (Johnston & Dover, 1987) and the Trp repressor (Otwinowski *et al.*, 1988). The DNA-binding constant of the SPKK motif must be much higher than that of the SPXX motif, which lacks the basic residues which would increase its DNA-binding ability. Therefore, Suzuki *et al.* (1990) proposed that while linker DNA is protected by SPKK-rich H1 arms, transcription factors can only slide along DNA rather than bind the same DNA sequences with their SPXX motifs. However, if the DNA binding ability of H1 SPKK motifs is weakened by phosphorylation of Ser residues, SPXX motifs of transcription factors could bind to linker DNA and thereby enter the nucleosome core to find the specific DNA sequence that will satisfy its specific DNA-binding region. This hypothesis has been indirectly supported by studies on the differential DNA-binding affinity of phosphorylated and non-phosphorylated sea urchin testis-specific H1 and H2B (Suzuki *et al.*, 1990; Green *et al.*, 1993).

1.2.9.2 Post-translational modification of linker histone

Linker histones are known to stabilise the higher order chromatin structure and are further involved in the salt-induced compaction of chromatin. It is reasonable to suggest that a charge neutralisation between histones and DNA must be involved in these processes. The most studied post-translational modification of linker histones is that of phosphorylation. Linker histones are phosphorylated at serine and threonine residues located in both the N- and C-terminal tail domains of the proteins. The C-terminal tail domain of linker histones is the primary site of many regulated phosphorylation events (Roth & Allis, 1992).

It has been proposed that phosphorylation of linker histones during the cell cycle should decrease the affinity of linker histones for DNA and lead to a loosening of the chromatin structure (Hall & Cole, 1986; Lea, 1987; Suzuki *et al.*, 1990; Aubert *et al.*, 1991; Hill *et al.*, 1991; Green *et al.*, 1993). In general, H1 is phosphorylated through S-phase, phosphorylation events that are probably associated with H1 deposition and chromatin replication (Muller *et al.*, 1985a). Through the G_2 phase, all H1 molecules are subject to increasing phosphorylation to reach a hyperphosphorylated state at metaphase. The level of phosphorylation can be as high as 22-24 phosphates per H1 molecule. Immediately following nuclear division, these hyperphosphorylated H1s are rapidly dephosphorylated to the S-phase levels (Bradbury *et al.*, 1974; Mueller *et al.*, 1985a). It has been suggested that the major mitotic kinase ($p34^{cdc2}$ kinase or maturation promoting factor (MPF)) is responsible for H1 phosphorylation in eukaryotic cells (Dunphy & Newport, 1988). H1 kinases or MPF control the phosphorylation through their recognition of the consensus motif (Ser/Thr-Pro-X-Lys/Arg) contained in the H1 molecule (Suzuki *et al.*, 1990; Davie, 1996).

The other modification to affect linker histones is poly-ADP ribosylation of H1(de Murcia *et al.*, 1986). As the internucleosomal DNA and bound histone H1 are hidden in the interior of the higher-order structure, it is believed that the competition

of the basic residues of histone H1 by their association with highly negatively charged poly-ADP ribose complexes may reduce their affinity for DNA and consequently affect the stability of higher order structure (Aubin *et al.*, 1983; de Murcia *et al.*, 1986; Yoon *et al.*, 1996).

1.2.10 30 nm chromatin structure

Studies of the arrangement of nucleosomes in the higher order chromatin fibre have been facilitated by virtue of the reversible salt-induced transition between the unfolded beaded chain of nucleosomes and the condensed 30-nm fibre (Thoma & Koller, 1977; Thoma *et al.*, 1979). Electron micrographs of chromatin undergoing this transition often show a fairly regular zig-zag pattern of chromatosome-linker DNA- chromatosome (Thoma *et al.*, 1979; Worcel *et al.*, 1981), but as compaction proceeds, the individual nucleosomes are no longer resolved and their arrangement in the 30-nm fibre becomes less accessible to direct observation. However, a substantial body of data concerning the biophysical and biochemical properties of both compact and relaxed chromatin has been accumulated, which places many constraints on their possible architecture (McGhee & Felsenfeld, 1980). On the basis of such data, there are three classes of helical models for chromatin fibres: the solenoid, twist-ribbon, and cross-linker models, and one nonhelical class, called the supranucleosomal particle model. Each model has common features that distinguish its members from those of other classes.

1.2.10.1 The solenoid models

Based on the assumption that nucleosomes are 11 nm diameter spherical particles, Finch & Klug (1976) postulated that a nucleosome chain can fold into a solenoid-like shape which is a single-start contact helix with a pitch of 11 nm and a diameter of about 30 nm, containing from 4 to 10 nucleosomes per turn, with a mode at about 6 or 7. However, the observation of a zigzag fibre at low ionic strength makes simple helical coiling unlikely, and Thoma *et al.* (1979) proposed an

alternative hypothesis concerning the condensation of chromatin fibres which interpreted the zigzag fibre as a contact helix with two nucleosomes per turn stabilised by H1-H1 interactions in the centre. As shown in Figure 1-5, condensation of the zigzag fibre into a compact fibre was postulated to occur by twisting this helix around its axis, increasing the number of nuclesomes per turn to 6 to 8, decreasing the angle between nucleosome disks while keeping the pitch constant at about 11 nm and maintaining the radial orientation of the nucleosomes. In the formation of higher order structures, the H1-binding regions on neighbouring nucleosomes come closer together so that an H1 polymer may be formed in the centre of the superhelical structure. Furthermore, solenoidal models can accommodate variable lengths of linker DNA which may be internally located provided the linker is allowed to form reversed loops, kinks or sharp bends (Butler 1984).

By using electric dichroism to investigate the orientation of nucleosomes in chicken erythrocyte chromatin in Mg^{++}-condensed 30 nm solenoids, McGhee *et al.* (1980) showed that the flat faces of chromatosome disk must lie close to parallel $(\pm 20)^{o}$ to the filament axis and suggested that the chromatosomes were arranged radially within the solenoid, with a maximum tilt angle of 44^{o} between the solenoid axis and the flat faces of the chromatosome discs. The most reasonable disposition for the spacer DNA within the 30 nm chromatin solenoid would be as a supercoil, wound around the helical path passing through the radially arranged chromatosome centres, and with the same diameter as the individual chromatosomes. With the further reasonable assumption that the interchromatosome spacer DNA is supercoiled, they estimated the average tilt angle for this DNA to be 20^{o}-25^{o}. Furthermore, McGhee *et al.* (1983) used electric dichroism to study the arrangement of nucleosomes in 30 nm chromatin solenoidal fibres prepared from a variety of sources (CHO cells, HeLa cells, rat liver, chicken erythrocytes and sea urchin sperm), and proposed the supercoiled spacer model which pointed out that the linker DNA,

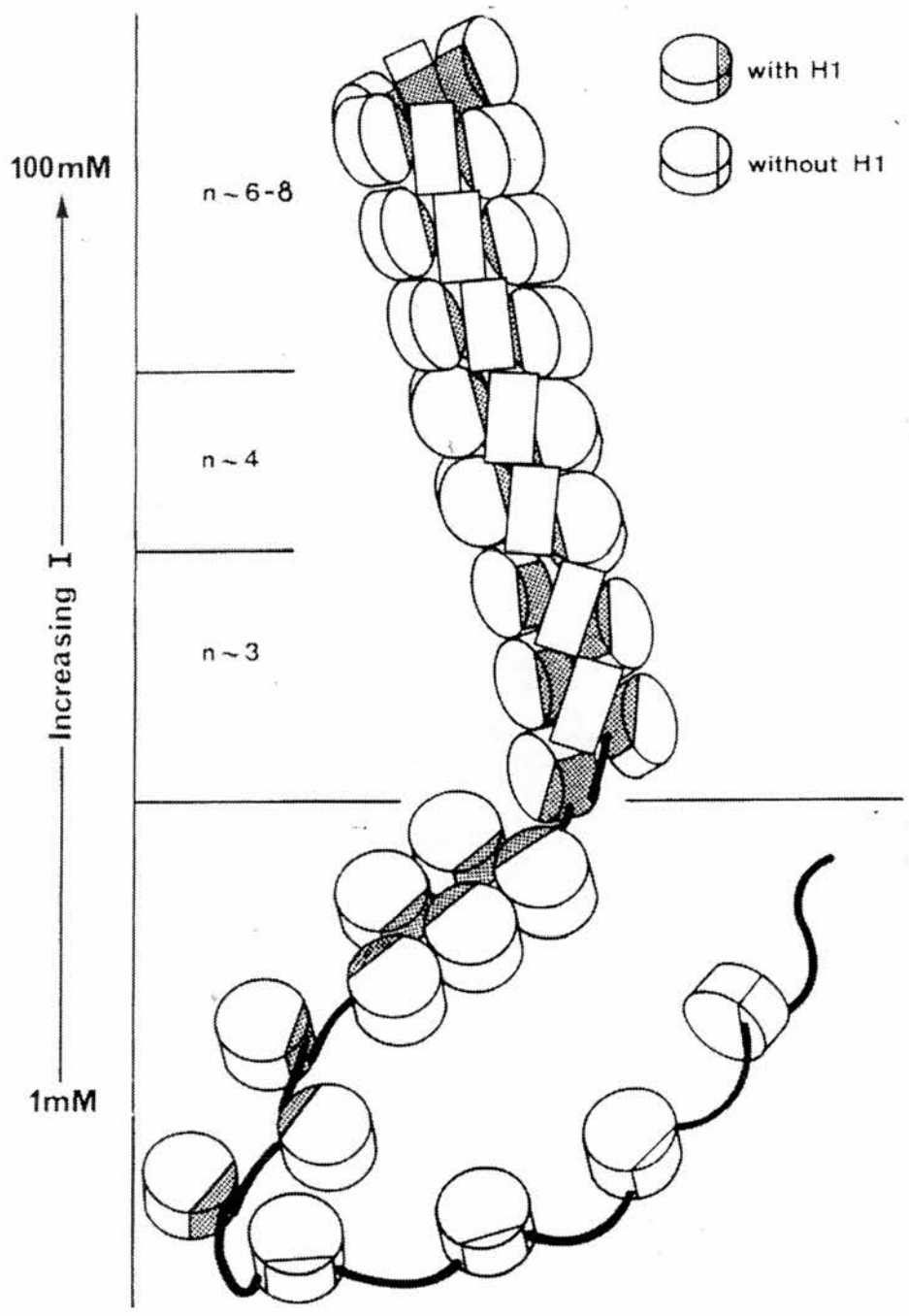


Figure 1-5 Helical superstructure formed by chromatin containing H1 with increasing ionic strength. The numbers of nucleosomes per turn are indicated by n. (Taken from Thoma *et al.*, 1979)

which varies with the sample chromatin studied, continues the superhelical path of the nucleosomal DNA with about 80 bp per turn. This leads to variable orientation of the DNA entry and exit sites of the nucleosomes in the fibre; nucleosomes connected by linkers of 40 bp would have the entry and exit sites alternating inside and outside, whereas H1 associated with 20 or 60 bp linkers would disrupt spacing between gyres of the solenoid. The supercoiled spacer model implies that the repeating unit of the solenoid superhelix could depend on spacer length, and the chromatin solenoid could have two distinguishable internal repeats, the number of nucleosomes per repeating unit and the number of nucleosomes per solenoid turn. Furthermore, if linker DNA is continuously supercoiled and H1 is in the centre of the fibre, H1 must be able to bind at many sites on the nucleosome, not only at the entry and exit point. However, the binding site of the lysine-rich histone, H1 or H5, has been placed at the point of exit and entry of the DNA strands onto the chromatosome surface (Simpson, 1978; Thoma *et al.*, 1979; Allan *et al.*, 1980), and from neutron scattering, Graziano *et al.* (1994) directly determined that the H1 is located in the interior of the filament.

1.2.10.2 The twisted-ribbon class of models

The basic supranucleosomal structure of chromatin is a zigzag helical ribbon with a repeat unit made of two nucleosomes connected by a relaxed spacer DNA(Thoma *et al.*, 1979; Worcel *et al.*, 1981). In consideration of linking numbers, Worcel *et al.* (1981) and Woodcock *et al.* (1984) proposed that the ribbon is wrapped on the surface of a cylinder to form a fibre with linker DNA that zigzags up and down the helix axis. Worcel *et al.* (1981) suggested that three different ribbon models could be constructed, each with a characteristic linking number increment (ΔL), which describes the change in the DNA linking number per nucleosome. He showed that the repeat unit of the proposed structure containing two nucleosomes with $-1^{3}/_{4}$ DNA turns per nucleosome and one spacer crossover per repeat, contributes -2 to the linking number of closed circular DNA. For a twisted ribbon structure, this meant that $\Delta L = -n$, where n= the number of nucleosomes.

Although the linking number of native SV40 minichromosomes and nucleosomes reconstituted onto circular DNA is -1 per nucleosome, this could be due to partial unwinding of nucleosomal DNA or a lack of contributions from H1 (McGhee & Felsenfeld, 1980; Wang, 1982). Therefore, Woodcock *et al.* (1984) suggested an alternative $\Delta L = -2$ for the helical ribbon structure. In their structure, the flat zigzag ribbon is coiled into a helix with a diameter of about 30 nm. The structure is independent of linker length and H1 molecules are neither completely exposed on the outside of the fibre nor completely hidden in the interior. Direct mass values for individual isolated fibres obtained from electron scattering measurements showed that the mass per unit length was dependent on ionic strength and ranged from 6.0 X 10^4 daltons/nm at 10 mM NaCl to 27 X 10^4 daltons/nm at 150 mM NaCl salt. These values are equivalent to 2.5 nucleosomes/11 nm and 11.6 nucleosomes/11 nm at the respective salt concentration. There are no apparent constraints on the handedness or the diameter of the helix formed, although Woodcock *et al.* (1984) maintained that the diameter should be independent of linker length. In addition, the model predicts that the fibre has a pitch of 32 nm and 18 nucleosomes per turn, and since the width of the ribbon depends on linker length, the model also predicts linker length-dependent variations in spacing along the fibre. However, this latter proposal is not supported by the X-ray diffraction studies carried out by Widom *et al.* (1985) and Widom & Klug, (1985).

1.2.10.3 The cross-linker models

The feature of these models is a nonsequential arrangements of nucleosomes cores connected by linker DNA which transverses the fibre axis. Furthermore, these models can be characterised by the number of helical ramps in each helical repeat. Based upon the tendency of nucleases to produce dinucleosomes rather than mononucleosomes, Staynov (1983) proposed two cross-linker models that form single start left-handed helices. Makarov *et al.* (1985) proposed a triple helix according to their flow dichroism measurements. Their model suggested (i) that nucleosomal disc faces are tilted relative to the fibre axis, (ii) that the orientation of nucleosomes does not change upon folding and unfolding of chromatin and (iii) that

the orientation of nucleosomes is maintained by the globular domain of histone H1. Williams *et al.* (1986) collected electron image and x-ray scattering data from nuclei and isolated chromatin fibres of seven different tissues, and suggested a linker length dependence of fibre diameter and mass per unit length. A structure with left-handed helical symmetry, a pitch of about 26 nm and a pitch angle of about 32° was calculated from three optical diffraction patterns of electron micrographs of negatively stained fibres. They proposed that the zigzag fibres form a two-start contact helix with the zig nucleosomes starting on one side and forming a ramp and the zag nucleosomes starting on the opposite side to form the second ramp. The linker DNA would cross the centre of the fibre. The model fits well with a constant, internal location of H1. However, for a very long linker DNA or no linker DNA, the model's structure may conflict with electric dichroism data or appear to be sterically impossible (Pederson *et al.*, 1986).

1.2.10.4 The supranucleosomal particle models

The mode of histone H1 binding to DNA undergoes a salt-dependent transition around 20 mM NaCl or 0.8 mM $MgCl_2$ (Renz & Day, 1976). Renz *et al.* (1977) studied the cooperative nature of the nucleosome-histone H1 interaction and suggested that above the transition point chromosome fibres appear to be composed of tandem arrays of globular structures. From electron microscopic and biochemical observations it has been proposed that the thick fibre is organised as a repeating series of granular 30-nm superbeads (Zentgraf & Franke, 1984).

It is important to consider the structure of the 30-nm filament, because it is the native form of the bulk of chromatin during interphase. The detailed structure of the chromatin fibre remains controversial and not yet well known. The solenoid model has gained substantial support from the studies of Widom and co-workers who used transmission electron microscopy and dynamic light scattering analyses to provide evidence that the linker DNA in dinucleosomes does in fact contract or fold as the salt concentration is raised and that the same changes occur when linker histones are removed (Yao *et al.*, 1990; Yao *et al.*, 1991). Furthermore, neutron

scattering data has shown that histone H1 is most likely to be located in the interior of the chromatin fibre (Graziano *et al.* 1994). However, in recent years, scanning-force microscopy and cryo-electron microscopy indicate that the 30-nm filament exists at low ionic strength as an irregular helix (Leuba *et al.*, 1994; Woodcock, 1994).

The key points in the absence of a precise model for the higher order chromatin structure lie in the lack of an identification for the path of the linker DNA as it passes from one nucleosome to its neighbours and the location of the linker histone within the fibre and the manner in which this protein determines and maintains the folded state.

1.2.11 The location of linker histone H1 in the higher order chromatin structure

The linker histone binds to the nucleosome and is essential for the organisation of nucleosomes into the 30-nm filament of chromatin. It is important to know how linker histones effect the folded state by virtue of their location in the higher order chromatin structure. From solid-phase immunoassay and inhibition experiments, based on accessibility to antibodies, results suggest that the globular domain of H5 is internally located in the 30 nm chromatin fibre (Dimitrov *et al.* 1987). Similarly, investigations by radioimmunoassay and electron microscopy indicated an inaccessible location for the folded domain of H5 in the higher-order chromatin fibre (Cattini *et al.*, 1988). However, other results suggest that the antigenic determinants of the globular domain of H5 are accessible to the antibody both in folded and unfolded chromatin, while those of the same region of H1 are masked in both states (Russanova *et al.* 1987). The use of proteolytic enzymes immobilised on membranes revealed that both GH1 and GH5 were inaccessible in folded chromatin fibres (Leuba *et al.* 1993). In summary, these investigations have yielded conflicting results.

Recently, Graziano *et al.* (1994) use neutron scattering to investigate chromatin structure. By altering the H_2O/D_2O ratio and deuterating H1, they were able to obtain information directly about H1 radial location in the filament. They found that chromatin 30-nm filaments have ~6 nucleosomes per 11 nm, with histone H1 at about the same radial location as the inner face of the nucleosome. This result together with the fact that H1 interacts with the nucleosome at the entry and exit points of nucleosomal DNA (Allan *et al.* 1980), is consistent with the model that H1 binds to the face of the nucleosome that is presented to the interior of the filament. Besides, the capacity of the globular domain of the chicken erythrocyte linker histone H5 to self-associate in solution (Maman *et al.* 1994) may also imply that linker histone globular domain occupies an axial position within the higher order chromatin fibre. The spatial juxtaposition of the GH5 domains at this location would be expected to promote their association and exert a stabilising effect upon higher order chromatin structure (Maman *et al.* 1994).

1.2.12 Higher order chromatin structure

In order to pack chromatin into a metaphase chromosome or indeed an interphase nucleus, it is necessary for the 30 nm chromatin fibre to be further folded or coiled. Many studies have shown that interphase nuclei and metaphase chromosomes have at least some global higher-order structure (Mathog *et al.*, 1984; Lichter *et al.*, 1988; Yakoto *et al.*, 1995). Intensive studies have suggested that mitotic chromosomes are formed by folding the 30 nm fibre into a series of helically arranged loops, each containing about 50 – 200 kbp of DNA. These loops may be further folded or condensed to achieve the packing found in the metaphase chromosome (Adolph *et al.*, 1977; Paulson & Laemmli, 1977; Callan, 1986). Therefore, the compaction of chromosomes may occur at least in part by the formation of an axial loop structure followed by further folding of the loops (Gasser & Laemmli, 1986). The folding of the loops is invariant (Boy de la Tour & Laemmli,

1988; Baumgartner *et al.*, 1991) and therefore the mitotic chromosome condensation occurs by a deterministic process.

Various factors involved both in promoting and maintaining higher order chromatin structure. For example, *in vivo* analysis of *Drosophila* chromosomes has shown that condensation is initiated at specific chromosome foci (Hiraoka *et al.*, 1989). Chromosome condensation may be directed by *cis*-acting DNA sites. Scaffold attachment regions (SARs), DNA sequences that bind with high affinity to proteins of the scaffold fraction, might also be an important component of the condensation process (Mirkovitch *et al.*, 1984; Mirkovitch *et al.*, 1988). Topoisomerase II (Uemura *et al.*, 1987; Adachi *et al.*, 1991; Hirano & Mitchison, 1993), SMC (Falk & Walker, 1988; Chuang *et al.*, 1994; Holt & May, 1996), and histones (Guo *et al.*, 1995; Bradbury, 1992) have been implicated in the process of condensation by means of their interaction with *cis* acting DNA elements.

As sister chromatids intertwine and become catenated during replication, topoisomerase II is required to facilitate the formation of proper folding under these conditions (Uemura *et al.*, 1987; Hirano & Mitchison, 1993; Swedlow *et al.*, 1993). Therefore, topoisomerase II is needed for the establishment but not the maintenance of condensation.

The function of the SMC family of proteins appears to be mainly restricted to mitotic chromosome condensation. In studying XCAP-C, an SMC member in *Xenopus laevis*, Hirano & Mitchison, (1994) have shown that chromosome condensation in vitro fails when anti-XCAP-C antibodies are present in assembly extracts. Also, previously assembled and condensed chromosomes can decondense slightly in response to treatment with the same antibodies, suggesting that XCAP-C plays an important role in the establishment and maintenance of the condensed chromosome.

Although linker histone H1 is an important factor in the formation of higher order chromatin structure, many studies have suggested that it is not required for chromosome condensation (Ohsumi *et al.*, 1993; Guo *et al.*, 1995; Shen *et al.*, 1995). Therefore, the role of linker histone H1 in this process remains to be clarified.

1.3 Controls of Transcription by Nucleosomes

It has been shown that most, if not all eukaryotic promoters and regulatory elements are organised into precise architectures within extensive nucleoprotein complexes. To activate transcription, the transcription factors have to compete with histones for DNA sites which may be blocked or obscured by nucleosomes on the promoter (Knezetic & Luse, 1986; Losa & Brown, 1987; Lorch *et al.*, 1987; Wu & Lichten, 1994; Grosschedl *et al.*, 1994). The wrapping of nucleosomal DNA in 1.75 superhelical turns over the surface of the histone octamer may be sterically incompatible with the assembly of a stable transcription initiation complex. Therefore, the issue about how nucleosome structure is destabilised in order to facilitate access to sequence-specific transcription factors and the general transcriptional machinery has been a focus of vigorous research in recent years.

Nucleosomes can be directed to precise positions on DNA by signals in the sequence (Simpson, 1991; Thoma, 1992), a phenomenon termed nucleosome positioning. Positioned nucleosomes can regulate the exposure of *cis*-acting DNA elements, preventing their interaction with *trans*-acting transcription factors and thereby inhibiting transcription. Nucleosome positioning signals may also play an important role in gene regulation by virtue of their proposed ability to direct the folding of a gene into an inactive higher order chromatin structure (Davey *et al.*, 1995). It follows that factors which modulate nucleosome positioning or the fundamental properties of chromatin will also have the potential to regulate gene activity. Similarly, the binding of linker histones to nucleosomes promoting the

formation of higher order chromatin structure will have the potential to influence transcriptional activation. However, how linker histones modulate gene activity in the context of chromatin structure is still a debatable issue (Caplan *et al.*, 1987; Kamakaka & Thomas, 1990). This is at least partly due to the lack of the information concerning the location and binding mode of linker histones in chromatin.

1.3.1 Interplay of transcription factor binding and nucleosome structure

Nucleosome perturbation during DNA replication might be required to provide a window of opportunity in which transcription factors access their recognition sequences prior to nucleosome reassembly (Wolffe, 1991). However, many studies have revealed that the binding of inducible transcriptional activators can initiate the formation of DNase I hypersensitive sites (DHSs). DHSs reflect an increase in sensitivity at specific DNA sequences (sites) to digestion by the nuclease such as DNase I. This structural perturbation correlates to an alteration of chromatin structure. DHSs occur primarily at promoter and enhancer regions prior to, or concurrently with, the induction of gene transcription (Gross & Garrard, 1988). Therefore, replication is not essential for remodeling chromatin structures of many genes (Tsukiyama *et al.*, 1994; Wall *et al.*, 1995; Hager *et al.*, 1995). Taken together these results indicate that there can be two types of hypersensitive site at a promoter: preset or remodellable. Preset hypersensitive sites are present before the activation signal is received; cell-type specific patterns of preset hypersensitive sites are generated during development. In contrast, the nucleosome arrays of remodellable hypersensitive sites change as part of the activation process. In many instances, the formation of DHSs, as a result of nucleosome remodelling, is directly linked to the binding of upstream activator proteins and is a prerequisite for transcriptional initiation. This step clearly represents an early stage in the gene activation pathway (Elgin, 1988; Hager *et al.*, 1995; Svaren & Horz, 1996; Svaren & Horz, 1997).

As both the presence or the absence of histones at DHSs has to be considered (McGhee *et al.*, 1981; Bresnick *et al.*, 1992; Truss *et al.*, 1995), a model has been proposed to describe alternative schemes by which transcriptional activators may initiate the formation of DHSs (Steger & Workman, 1996). Transcriptional activators may bind to their recognition elements in nucleosomal DNA, resulting in the formation of a ternary complex containing activators, histones and DNA. As a consequence of binding, the underlying histone-DNA contacts are partially disrupted, which may result in histone displacement from the transcription factor-bound sequences. The nucleosome-free gap generated would certainly represent a DHS, since the loss of histones would expose the DNA surrounding the bound factors to DNase I. Alternatively, the co-binding of activators with histones in a ternary complex could sufficiently perturb local nucleosome structure so as to increase its DNase I sensitivity.

Many transcription factors have an ability to compete with histones for access to binding sites in nucleosomes although they vary in their affinity for nucleosomal DNA. These include glucocorticoid receptor (Perlmann & Wrange, 1988; Pina *et al.*, 1990; Archer *et al.*, 1991), GAL4 (Vettese-Dadey *et al.*, 1994), Sp1 (Li *et al.*, 1994; Chen *et al.*, 1994), USF (Adam & Workman, 1993; Chen *et al.*, 1994), NF-κB (Adam & Workman, 1993), Myc/Max and Max/Max dimers (Wechsler *et al.*, 1994), TBP (Imbalzano *et al.*, 1994) and TFIII A (Hayes & Wolffe, 1992b; Lee *et al.*, 1993).

Nucleosome positioning will determine the location of transcription factor-binding sites incorporated within the nucleosome structure (Vettese-Dadey *et al.*, 1994; Adams & Workman, 1995). Polach & Widom (1995) have proposed that DNA sequences may be transiently released from the surface of the histone octamer by partial unpeeling of the DNA ends from the nucleosome edge. The probability of a sequence being transiently exposed is therefore greater near the edge of the nucleosome than at its centre. Studies with mononucleosomes bearing five GAL4 sites have shown that GAL4 binding initially occurs at the site close to the edge of the nucleosome, followed by binding to the more internally located sites (Vettese-

Dadey *et al.*, 1994; Adams & Workman, 1995). The results also showed that the initial binding of GAL4 to the edge of the nucleosome facilitates further factor binding to lower affinity sites.

Many studies have suggested that activator binding sites are frequently in clusters and that activators bind to these sites cooperativily (Ptashne, 1986; Taylor *et al.*, 1991; Oliviero & Struhl, 1991; Vettese-Dadey *et al.*, 1994). Therefore, the initial binding of an activator to a nucleosome creates a localized perturbation of histone-DNA interaction. This localized disruption enhances the nucleosome-binding affinity of factors for sites in the immediate vicinity, and their subsequent binding in turn breaks additional histone-DNA contacts, leading to a cascade of factor loading onto the nucleosome even at the most difficult positions (Adams & Workman, 1995). In addition, cooperatively in binding between GAL4-AH, USF and NF-κB in any combination demonstrates that the binding of one factor can stimulate the binding of additional factors to adjacent sites within a nucleosome core (Adams & Workman, 1993). This means that the disruption occuring when one factor binds a nucleosome can potentiate the binding of another factor that would otherwise bind its site only poorly. At the level of nucleosomal arrays, one or more of the binding sites must be located near the edge of a nucleosome. Because of the relatively weak interaction between DNA and the histone octamer at the nucleosome edge, productive binding of externally positioned sequence elements by their cognate factors is achieved. Therefore, through cooperative effects, this binding greatly stimulates occupancy of the neighbouring sites, which in turn greatly enhances the transcriptional level (Xu *et al.*, 1995; Steger & Workman, 1996).

1.3.2 Histone acetylation and gene activation

Numerous experiments have indicated that transcription of some genes is influenced by the acetylation of particular lysine residues in specific histones (Kayne *et al.*, 1988; Durrin *et al.*, 1991; Thompson *et al.*, 1994), and, conversely, that the

transcriptional silencing of specific loci is associated with reduced nucleosomal acetylation (Braunstein *et al.*, 1993). Although the exact relationship between histone acetylation and transcription is, as yet, undefined, the conserved tails of the core histones play a crucial role in this relationship (Lee *et al.*, 1993). These amino-terminal tail domains are lysine-rich and are the sole targets for acetylation. Acetylation greatly reduces the affinity of the histone H4 tail for DNA (Hong *et al.*, 1993). The physical consequences for nucleosomal integrity of acetylating all of the histone tails in the absence of any other proteins are relatively minor. However, there is a modest reduction in the wrapping of DNA around the histone octamer and nucleosomes pack together less efficiently in arrays (Norton *et al.*, 1989; Garcia-Ramirez *et al.*, 1995).

From the study of transcription factor TFIIIA binding to a positioned nucleosome on 5S rDNA, it was shown that the factor did not bind efficiently to 5S rRNA gene within a nucleosome if the core histones were not acetylated, but it did bind following acetylation of histones (Lee *et al.*, 1993). Likewise, GAL4 binding to nucleosomal DNA is also facilitated by acetylation of histone H4 (Vettese-Dadey *et al.*, 1996). These results suggest that acetylation of the amino-terminal tails substantially weakens the constraints on DNA imposed by the core histones. Therefore, histone acetylation may provide a molecular mechanism by which DNA can be rendered more accessible to trans-acting factors while still maintaining a nucleosomal architecture (Wade *et al.*, 1997).

Studies in yeast have demonstrated that Gcn5p (HAT A), a known yeast transcriptional coactivator, exists in a heteromeric complex with at least two additional partners, Ada2p and Ada3p (Marcus *et al.*, 1994; Horiuchi *et al.*, 1995). A combination of genetic and biochemical evidence suggest that this complex interacts with enhancer-binding factors and may establish a bridge between the upstream activating sequence and the basal transcription machinery in yeast (Marcus *et al.*, 1994; Horiuchi *et al.*, 1995; Georgakopoulos *et al.*, 1995; Silverman *et al.*, 1995; Barlev *et al.*, 1995). The Gcn5p component of the complex has the capacity to

acetylate specific lysine residues in histones H3 and H4, which are known to be associated with transcriptional activity (Kuo *et al.*, 1996). Similar enzymatic activities have been found for P/CAF (Yang *et al.*, 1996), p300/CBP (Ogryzko *et al.*, 1996; Bannister & Kouzarides, 1996), and $TAF_{II}230/250$ (Reese *et al.*, 1994; Ruppert & Tjian, 1995; Moqtaderi *et al.*, 1996; Sauer *et al.*, 1996). Although Gcn5p and P/CAF are related proteins, there is no significant sequence identity or known structural similarity with p300/CBP or $TAF_{II}230/250$. Thus, diverse proteins possess histone acetyltransferase activity.

As transcriptional coactivators can be histone acetyltransferases and other transcriptional regulators can deacetylate histones (Taunton *et al.*, 1996), the recruitment of coactivators could direct the local destabilization/stabilization of repressive histone-DNA interactions and transcription might be continuously controlled. In this scenario, repressive nucleosomes would prevent either the association or functioning of the basal transcriptional machinery on a particular promoter. Targeted acetylation would provide a means of allowing the basal transcriptional machinery to displace nucleosomes and to assemble a functional transcriptional complex. As core histone proteins remain associated with DNA in the vicinity of a promoter despite the recruitment of the basal transcriptional machinery (Nacheva *et al.*, 1989), the targeted or general activity of histone deacetylase will tend to return nucleosomes to their repressive configuration. The maintenance of gene activity would therefore require the continued activity of the coactivators as acetyltransferase. In this way, transcriptional activity could be continually modulated through variation in chromatin conformation. These results suggest that the eukaryotic transcriptional machinery is not only adapted to function in a chromatin environment, but has the potential to make use of the packaging of DNA to regulate genes (Wade *et al.*, 1997).

1.3.3 ATP-dependent nucleosome remodeling factors

SWI/SNF and NURF are chromatin remodeling factors which act to oppose the contacts between nucleosomal histones and DNA.

The SWI/SNF multiprotein complex is a strong candidate for a chromatin remodeling activity. Mutations in the SWI/SNF genes are supressed by mutations in genes coding for the nucleosome core histones and an HMG1-like protein, SIN1/SPT 2, suggesting that the SWI/SNF complex acts, either indirectly or directly, to antagonize the repressive association of DNA with chromosomal proteins (Kruger & Herskowitz, 1991; Hirshhorn *et al.*, 1992; Kruger *et al.*, 1995). A good example of the hypothesis that SWI/SNF stimulates transcription is the finding that the chromatin structure of the SUC2 promoter *in vivo* is more resistant to digestion with micrococcal nuclease in a *swi2/snf2* or *snf5* mutant, indicating a failure to antagonize nucleosomal organisation at the promoter region (Hirshhorn *et al.*, 1992; Matallana *et al.*, 1992).

Biochemical analysis of yeast and mammalian SWI/SNF complexes have demonstrated that the complex utilises the energy released from ATP hydrolysis to disrupt nucleosomal structure, which consequently stimulates the binding of transcription factors including TBP (Cote *et al.*, 1994; Kwon *et al.*, 1994; Imbalzano *et al.*, 1994). As revealed by DNaseI hypersensitive site analysis, purified SWI/SNF can also disrupt an array of preassembled nucleosomes reconstituted with purified histones in an ATP-dependent manner (Owen-Hughes *et al.*, 1996). The ATP-dependence of SWI/SNF chromatin disruption is due to the SWI2/SNF2 subunit, which contains an ATPase domain and has been shown to catalyze ATP hydrolysis in a DNA-dependent manner (Laurent *et al.*, 1993; Cote *et al.*, 1994). The SWI/SNF complex binds to naked DNA in an ATP-independent manner. It can induce positive supercoiling of plasmid DNA and has a preferential affinity in the nanomolar range for four-way junction DNA, a structure that may mimic the cross-over point where DNA enters and exits the nucleosome (Quinn *et al.*, 1996). Furthermore, the

SWI/SNF complex reduces in an ATP-dependent reaction the number of stable negative supercoils contained in nucleosome-assembled plasmid DNAs (Kwon *et al.*, 1994; Wilson *et al.*, 1996). Therefore, it is possible that SWI/SNF induces a conformational change in nucleosomal DNA, which alters nucleosome architecture. The energy released from ATP hydrolysis is required to modify the DNA-histone contacts (Steger & Workman, 1996).

NURF activity was first observed in a cell-free *Drosophila* embryo extract as being capable of acting in combination with the GAGA transcription factor to disrupt nucleosomes preassembled with a crude *Drosophila* chromatin assembly extract (Becker & Wu, 1992; Tsukiyama *et al.*, 1994; Tsukiyama & Wu, 1995). The chromatin remodelling properties of NURF have proved thus far to be very similar to those of SWI/SNF (Cote *et al.*, 1994; Tsukiyama & Wu, 1995; Owen-Hughes *et al.*, 1996). However, the physical properties of NURF clearly distinguish it from SWI/SNF. NURF is considerably smaller, being composed of only four protein subunits in contrast to at least 10 subunits in SWI/SNF complex. The ATPase activity of NURF is stimulated more substantially by the presence of nucleosomes than by free DNA or core histone alone, suggesting that NURF is capable of detecting some unique feature of nucleosome organisation. This difference and the observation that only substoichiometric levels of NURF are needed to facilitate GAGA-factor-mediated nucleosome perturbation – one NURF per 18 nucleosomes – suggests a mechanism of action that is distinct from SWI/SNF (Tsykiyama & Wu, 1995). Although one of the NURF subunit, ISWI, has significant identity to SWI2/SNF2 over the ATPase domain (Elfringe *et al.*, 1994), they appear to comprise two distinct classes of ATP-dependent chromatin remodelling activities.

Once positioned at a particular chromosomal domain, SWI/SNF or NURF may participate in the sequence-specific remodeling of chromatin observed for numerous genes *in vivo*. Although the disruption of nucleosomes by either SWI/SNF or NURF occurs independently of bound transcription factors, this disruption is necessary for enhanced binding of transcriptional activators to nucleosomal DNA

(Cote *et al.*, 1994; Kwon *et al.*, 1994; Imbalzano *et al.*, 1994;Tsukiyama & Wu, 1995). Thus, while SWI/SNF and NURF interact with nucleosomes in a non-sequence-specific manner, the induced binding of the transcription factors is sequence-specific. SWI/SNF or NURF-stimulated factor binding, therefore, could result in increased site-specific nuclease sensitivity (for example, DHSs) as a consequence of the bound factors (Owen-Hughes *et al.*, 1996).

1.3.4 Mechanisms of nucleosome remodelling

It is well established that changes in chromatin structure accompany gene activation or potentiation to a pre-activated state (Elgin, 1988; Gross & Garrard, 1988). Whether nucleosomes are lost or reconfigured during this process has yet to be clarified (Thoma, 1991; Hayes & Wolffe, 1992b; Adams & Workman, 1993; Paranjape *et al.*, 1994).

Histone octamers are capable of relocating on a DNA fragment in *cis* (Beard, 1978) and possess an inherent localised mobility that can lead to repositioning over short distances (Pennings *et al.*, 1991). Reconfiguring chromatin structure at gene regulatory domains by nucleosome sliding would require multiple nucleosomes to be moved to new positions. Recent studies have suggested that entire arrays of nucleosomes can be rendered mobile by an ATP-driven process (Wall *et al.*, 1995; Varga-Weisz *et al.*, 1995). Furthermore, the ATP-dependent activities stimulated binding of GAGA factor to a heat shock gene promoter, and even enhanced the accessibility to restriction enzymes for the nucleosome-assembled plasmid DNA. The same activities may have been responsible for transcription factor-induced chromatin remodeling in *Drosophila* embryo extracts (Pazin *et al.*, 1994). These results indicate that chromatin is very dynamic, and that ATP-dependent processes constantly mobilise nucleosomes to facilitate transcription factor access. On the basis of these results, they proposed that binding of factors to the chromatin template occurs first, followed by an ATP-dependent reconfiguration of nucleosome structure.

Finally, binding of activator to transcriptional activation domain facilitates the transcription process (Pazin et al., 1994).

The complete dissociation of histones from the DNA as means to generate a nucleosome-free region upon transcription factor binding is another possible pathway for nucleosome remodelling. As histone dissociation would require acceptor molecules to which the histone can transfer (Workman & Kingston, 1992), this method of histone dissociation would depend upon the nature of the histone acceptor.

In vitro studies with two different nucleosome assembly proteins, nucleoplasmin and nucleosome assembly protein 1 (NAP-1), demonstrated that these proteins can facilitate transcription factor binding and nucleosome disassembly by acting as histone chaperones (Owen-Hughes & Workman, 1996). Moreover, when multiple factors are specifically bound to a nucleosome within an array of nucleosomes, nucleoplasmin facilitates disassembly of the histone octamer specifically in the factor-bound nucleosome (Owen-Hughes *et al.*, 1996). The H2A/H2B dimers are removed first, followed by loss of the H3/H4 tetramer (Chen *et al.*, 1994). When H2A/H2B dimers are covalently protein-protein crosslinked to the H3/H4 tetramer so that the histone octamer can not come apart, histone displacement by nucleoplasmin or NAP-1 is inhibited (Walter *et al.*, 1995). Taken together, these results suggest that the nucleosome assembly proteins, nucleoplasmin and NAP-1 may participate in chromatin remodeling by disassembling nucleosomes bound by transcription factors.

An alternative pathway of histone dissociation is octamer transfer, a process which involves relocating an intact histone octamer onto another fragment of DNA. Evidence has shown that a histone octamer can relocate to the rear of an advancing polymerase, via an intermediate state where the octamer is simultaneously associated with DNA both in front of, and behind, the polymerase (Studitsky *et al.*, 1995). The most concentrated DNA surrounding a factor-bound nucleosome is in the adjacent regions within the same nucleosomal array. It is therefore likely, that octamer

transfer would most often place the histone octamer at another location on the same region of chromosomal DNA. Thus, histone octamer transfer may also contribute to nucleosome relocation in *cis* (Steger & Workman, 1996).

1.3.5 The role of HMG proteins in the regulation of gene expression

The HMG (high mobility group) chromosomal proteins are among the most abundant and ubiquitous non-histone proteins. The HMG proteins are intimately associated with chromatin, may influence the structure of the chromatin fibre and affect the binding of transcriptional regulatory factors to their targets. Due to lack of a functional definition, the HMG proteins are currently classified by their chemical and physical properties based on the properties of calf thymus HMG proteins. The HMG proteins include three major families HMG 1/2, HMG 14/17 and HMG I/Y, which have different functions in various cellular processes such as replication, transcription or nucleosome assembly.

HMG 1/2 proteins can bind to DNA through their HMG box, a highly charged C-terminal region flanking the folded domain (Wisniewski & Schulze, 1994). This binding can stimulate transcription by facilitating the formation of an active initiation complex on template DNA (Tremethick & Molloy, 1988). It has been found that the HMG box is necessary for this (Aizawa *et al.*, 1994). In addition, Varga-Weisz *et al.* (1994) have also suggested that HMG-1 can replace histone H1 from the nucleosome to activate transcription.

The HMG 14/17 proteins are the only chromosomal proteins known to have a higher affinity for nucleosomal cores than for DNA (Albright *et al.*, 1980; Sandeen *et al.*, 1980). It has been suggested that the location of HMG 14/17 proteins within the chromatosome may overlap the binding sites of linker histones and therefore affect linker histone binding to nucleosomal DNA (Alfonso *et al.*, 1994).

The HMG I/Y proteins have been shown to be required for the transcription activities of both NF-κB and ATF-2 in the IFN-β gene promoter by virtue of their binding to the minor groove of DNA (Thanos & Maniatis, 1992; Du *et al.*, 1993). It has been proved that HMG I/Y proteins prefer to bind to A/T rich DNA, induce positive supercoiling and therefore unwind DNA (Reeves & Nissen, 1993; Nissen & Reeves, 1995). To sum up, the HMG I/Y proteins may influence the positioning of nucleosomes by virtue of their ability to bind preferentially to A/T rich DNA and displace H1 effectively.

Although the HMG proteins seem to be involved in gene regulation through their interaction with nucleosomes and DNA, their biological function is still largely speculative and needs to be further identified.

1.3.6 The role of linker histone in the regulation of gene expression

A role for linker histones as repressive components of positioned nucleosomes has been suggested by a series of biochemical experiments using the mouse mammary tumor virus (MMTV) long terminal repeat (LTR) and the *Xenopus* 5S rRNA genes (Richard-Foy & Hager, 1987; Perlmann & Wrange, 1988; Pina *et al.*, 1990; Bouvet *et al.*, 1994; Kandolf, 1994). These investigations indicated that removal of histone H1 might be sufficient to allow the transcriptional machinery to function effectively in chromatin. Furthermore, the mobility of histone octamers positioned on constructs of 5S rRNA genes is suppressed by the binding of histone H1/H5 (Pennings *et al.*, 1994). Mobile histone octamers on the 5S rRNA genes are accessible to the pol III transcriptional machinery, but in the presence of linker histone H1 this accessibility is severely restricted and transcription is repressed (Ura *et al.*, 1995; Ura *et al.*, 1996). The removal of linker histone H1 from the chromatin of the MMTV LTR, when the glucocorticoid receptor initiates the process of chromatin rearrangement, might increase the local mobility of histone octamer with respect to DNA sequence, facilitating transcription factor access to regulatory

elements (Bresnick *et al.*, 1992). Therefore, the reversible restraint on chromatin dynamics may play a role in local regulation of processes that require access to the DNA (Pennings *et al.*, 1994) and the addition of histone H1 may restrict nucleosomes to a certain subset of possible translational positions (Ura *et al.*, 1995). Moreover, comparison of the structure of the globular domain of histone H5 with hepatocyte nuclear factor (HNF)-3γ, which is a mammalian tissue-specific transcription factor, has shown that their structure, DNA binding mode and conformation are similar (Ramakishnan *et al.*, 1993; Goytisolo *et al.*, 1996). This structure similarity may be functionally significant in respect of their effects on transcriptional control.

As the linker histone binds to DNA much less tightly than the core histones, it is the easier component of chromatin to remove. The removal of histone H1 from chromatin might involve active displacement by molecular machines such as RNA polymerase (Cote *et al.*, 1994; Wilson *et al.*, 1996) or the SWI/SNF complex (Cote *et al.*, 1994; Imbalzabo *et al.*, 1994). Alternatively, transcription factors might simply compete for their recognition sites in the nucleosome and displace H1 by virtue of their greater binding constant to DNA. Indeed, the binding of histone H1 to the nucleosome is very dynamic and can be redistributed (Caron & Thomas, 1981; Thomas & Rees, 1983). Therefore, the displacement of histone H1 by transcription factors is possible. In addition, this competition might also be influenced by the post-translational modification of histone H1 or by the presence of molecular chaperones. Many studies have shown that phosphorylation of linker histones weakens the interaction of linker histones with DNA *in vitro* (Hill *et al.*, 1991) and with chromatin *in vivo* (Aubert *et al.*, 1991), and phosphorylated linker histones are preferentially associated with transcriptionally competent chromatin (Roth *et al.*, 1988). Therefore, linker histone phosphorylation might also be a targeted function of the transcriptional machinery. As the molecular chaperone, nucleoplasmin, could direct the selective dissociation of linker histones from the chromatin of somatic nuclei (Dimitrov & Wolffe, 1996) and provide a binding site that is favoured compared to nucleosomal DNA, mechanisms that could direct the selective

recruitment of nucleoplasmin-like protein to a promoter might facilitate the removal of linker histones and therefore activate transcription (Wolffe *et al.*, 1997).

Recent studies have challenged the idea that H1 is essential for higher order structure of eukaryotic chromatin. The protozoan *Tetrahymena thermophila* contains two types of nuclei, a large transcriptionally active macronucleus and a small transcriptionally inert micronucleus. When the gene encoding either the macronuclear-specific or micronuclear-specific H1 variant is deleted, no effect is observed on viability or growth rate and the transcription of most genes does not change (Shen *et al.*, 1995; Shen *et al.*, 1996). Although the nucleus lacking H1 does increase in size, it appears that H1 plays a non-essential role in chromatin compaction and only controls the expression of a subset of genes (Shen *et al.*, 1995). Furthermore, in cell-free extracts of *Xenopus* eggs, linker histones were found to be nonessential for the assembly of nuclei which were deemed normal by cytological criteria and their ability to replicate DNA and to import proteins into the nucleus (Dasso *et al.*, 1994). Therefore, it remains to be determined whether loss of linker histones has more subtle rather than global effects on transcriptional regulation.

1.4 Thesis Perspective

The manner in which linker histone globular domains associate with the nucleosome is a key issue in understanding the location of the linker histone tails, and thus their functional roles in regulating gene expression (Crane-Robinson, 1997). However, studies on the binding of the linker histone globular domain to the nucleosome have not yielded a unified picture (Allan *et al.*, 1980; Hayes & Wolffe, 1993; Travers & Muyldermans, 1996). The main difference between these studies concerns the precise location of the globular domain with respect to the nucleosome dyad axis (Figure 1-4). Understanding of the location of the linker histone globular domain within the nucleosome can provide further insight into the nucleosome

architecture both in the context of the location relative to the dyad on a statistically significant nucleosome and possible association with different nucleosomes. Therefore, it would contribute to the implication of gene regulation and higher-order structure models.

Previous work in our laboratory led to the development of a technique for mapping the precise translational positions adopted by core histone octamers reconstituted onto long fragments of DNA. The approach was employed to study the chicken β-globin gene region and to provide the first long-range positioning map for an entire, contiguous gene and its flanking sequence (Davey *et al.*, 1995). The high resolution of the mapping approach and its capacity to provide quantitative information about nucleosome positioning sites makes it a useful approach for comparative studies. Using this mapping technique, I was able to generate *in vitro* nucleosome positioning maps to compare chromatosome and core particle positioning sites on the chicken adult β-globin gene region. By this approach, I was able to provide further insight into the binding mode of linker histone globular domain on the nucleosome, both in a local (location on individual nucleosome) and long-range (association with particular nucleosome) context.

Chapter 2 General Materials and Methods

2.1 Reagents and stock solutions

Acrylamide - 30% (W/V) acrylamide (29:1, acrylamide : N,N'-methylenebisacrylamide)stock solution (Anachem) for 15% protein gel electrophoresis. 40% (W/V) acrylamide (19:1, acrylamide : N,N'-methylenebisacrylamide) stock solution (Anachem) for 4, 6 and 8% DNA sequencing gel electrophoresis.

Agarose - SeaKem GTG agarose (FMC) or NuSieve agarose (FMC) was used for separation of DNA (≥1 kb or ≤1 kb, respectively).

Ammonium acetate – 5M ammonium acetate (pH 7.5) for neutralization of alkali-denatured DNA.

Ammonium persulphate - 10% (W/V) and 25% (W/V) fresh solutions of ammonium persulphate were made in water and used for polymerizing polyacrylamide gels.

Antibiotics - 100 mg/ml ampicillin (100X) and 50 mg/ml kanamycin (715X) stocks were made in distilled water and sterilized by filtration (0.45 μm). The solutions were stored at -20°C.

BSA (Bovine serum albumin) – A 10 % solution was prepared by dissolving 1 g BSA (Sigma) in 10 ml distilled water and then stored at –20 °C.

Dextran sulphate - 9% (W/V) dextran sulphate (Pharmacia) was used in filter hybridization as molecular crowding reagent.

DEPC-treated water – 0.1% of diethyl pyrocarbonate (Sigma) was added to distilled water for at least 12 hours and then autoclaved it.

Dialysis tubing - Dialysis membrane (MW 3500; Spectrum) was prepared by boiling in 1mM EDTA (pH 7.5) and 2 % $NaHCO_3$, washing with distilled water three times, and then adding 1mM EDTA. The tubing was autoclaved and stored at 4°C. This tubing was used for electroelution.

DNA Polymerase I large (Klenow) fragment - The Klenow fragment (Amersham) which lacks the 5′→3′ exonuclease activity of intact DNA Polymerase I but retains the 5′→3′ polymerase, the 3′→5′ exonuclease and the strand displacement activities was used in the monomer extension experiments.

DTT - To prepare a 1M solution, 3.09 g dithiothreitol (Sigma) was dissolved in 10mM Sodium acetate (pH 5.2) and sterilized by filtration, aliquotted and stored at -20°C.

EDTA - 250 mM solution was prepared by dissolving 93.5 g ethylenediaminetetraacetic acid (BDH) in a final volume of 1000 ml distilled water, the pH was adjusted to 8.0 with NaOH, and then sterilized by autoclaving.

EtBr (Ethidium Bromide) - 10 mg/ml solution was prepared by dissolving 1 g EtBr (BCL) in 100 ml distilled water and the solution was stored in a light-proof bottle at room temperature.

Grinding buffer – 50 mM Tris-Cl (pH 7.9), 2 mM EDTA, 0.1 mM DTT, 1mM β-mercaptoethanol, 0.5 M NaCl, 0.1 % sodium deoxycholate, 5 % glycerol and 25 μg/ml PMSF.

IPTG – 840 mM stock solution of isopropyl β-$_D$-thiogalacto-pyranoside (Sigma) was made in autoclaved water and stored at –20°C.

LB (L broth) - Luria-Bertani medium was prepared by dissolving 10 g bacto-tryptone (Difco), 5 g bacto-yeast extract (Difco) and 10 g NaCl in 1 L distilled water and autoclaving for 20 minutes at 15 lb/inch^2.

LB agar - 1.5g of agar was added to 100 ml LB medium and autoclaved as above. ampicillin was added to 0.1 mg/ml immediately before pouring plates.

Lysis buffer - 60 mM NH_4Cl, 60 mM KCl, 20 mM Tris-HCl (pH 8), 6 mM $MgCl_2$, 6 mM β-mercaptoethanol and 16% sucrose was used for ribosomal RNA preparation.

Micrococcal nuclease (MNase) - Enzyme (Worthington) which catalyzes the hydrolysis of deoxyribonucleic acid or ribonucleic acid to produce 3′- nucleoside phosphates was used for monomer DNA preparation.

Nitrocellulose filters (Hybond-N) - Nylon membrane (Amersham) used in Northern transfer experiments.

Phenol – AR grade solid phenol (Fluka) was melted at 68 °C and then 8-hydroxyquinoline was added to a final concentration of 0.1%. The melted phenol was extracted several times, firstly with 1 M Tris-Cl (pH 8.0) and then with 10 X TE until the phenol was saturated and the pH of the aqueous phase was greater than 7.6. The phenol was then frozen at –20 °C.

Phosphate-buffered saline (PBS) - 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 1.4 mM KH_2PO_4.

PMSF - Phenylmethylsulphonyl fluoride (Sigma) was prepared as 250 mM stock in propan-2-ol and stored at 37°C.

Potassium Acetate – To prepare 3M K/5M Ac, 29.5 g potassium acetate and 11.5 ml of acetic acid was made up to 100 ml with H_2O and was autoclaved.

Proteinase K – 20 mg/ml concentration of this enzyme (BDH) was prepared in distilled water and stored at –20 °C. This enzyme was used for alkaline phosphatase inactivation.

RNase A – Bovine pancreatic Ribonuclease A (BCL) was dissolved at 10 mg/ml in 10 mM Tris-Cl (pH 7.5) and 15 mM NaCl. The solution was boiled for 5 minutes to destroy any contaminating DNase I activity. The solution was aliquotted and stored at – 20 °C.

Salmon sperm DNA - 10 mg/ml stock was made by dissolving 500 mg salmon sperm DNA in 50 ml TE and placed on a roller overnight at 4°C to obtain a homogenous viscous solution. Sonication was applied to generate sheared salmon sperm DNA. After sonication, the solution was aliquotted and stored at -20°C.

SDS - 10% solution was prepared by dissolving 100 g Sodium dodecyl sulphate (Sigma) in 1 L distilled water.

Sephadex G-25/G-50 - Solid Sephadex G-25/G-50 (medium) was suspended in TE, left at room temperature over night, and then autoclaved for 15 minutes.

Sodium acetate - 3 M solution prepared by dissolving 408.1 g sodium acetate (BDH) in 1 L distilled water, and pH was adjusted to 5.2 with glacial acetic acid.

SSC - 1 L of 20X solution was made by dissolving 175.3 g NaCl and 88.2 g sodium citrate in DEPC-treated water. The pH was adjusted to 7.0 with NaOH, and then autoclaved.

Sucrose gradient solutions – 10% and 30% sucrose solutions containing 60 mM NH_4Cl, 60 mM KCl, 50 mM Tris-HCl (pH 8), 12 mM $MgCl_2$ and 6 mM β-mercaptoethanol were used for sucrose gradient preparation.

Taq DNA polymerase – Taq DNA polymerase (Pharmacia) was used to amplify specific sequences of DNA by PCR.

TBE - 20X solution was prepared by dissolving 121 g Tris, 62 g Boric acid, and 7.4 g EDTA in 1 L distilled water. The solution was autoclaved.

TE - 10 mM Tris and 0.1 mM EDTA prepared as a 50X solution.

Tri Reagent – Tri Reagent (Sigma) which is a mixture of guanidine thiocyanate and phenol in a mono-phase solution was used in the isolation of single-stranded DNA.

Tris-Cl – 1 M solutions were prepared at various pHs by titration with HCl in distilled water and sterilised by autoclaving.

2X YT medium - 2X YT medium was prepared by dissolving 16 g bacto-tryptone, 10 g yeast extract and 5 g NaCl in 1 L distilled water.

2.2 Generation of a recombinant plasmid DNA

2.2.1 Preparation of insert and vector DNAs

Purified DNA fragments we prepared from plasmid by digestion with appropriate restriction enzymes under manufacturers' indications and with the addition of 100 μg/ml BSA.

2.2.2 Purification of DNA from agarose gels

Restriction enzyme-treated DNA was separated by agarose gel electrophoresis, and the desired DNA fragments were collected by electroelution. The required band was excised from the gel, transferred to dialysis tubing and subjected to electrophoresis in 0.5X TE at a constant voltage of 200 V for 25 minutes. The DNA-containing solution in the dialysis tubing was collected and purified by phenol and chloroform/isoamyl alcohol extraction and ethanol precipitation.

2.2.3 Dephosphorylation of DNA

5′ terminal phosphates of DNA fragments were removed with calf intestinal alkaline phosphatase (CIP). 1 unit of this enzyme (Boehringer Mannheim) was added to DNA in 50 mM Tris-HCl and 0.1 mM EDTA, pH 8.5 and incubated 37 °C for 30 minutes. Another 1 unit of alkaline phosphatase was then added for a further 30 minutes incubation. The sample was then made 5 mM EDTA, 0.5% SDS and 50 μg/ml Proteinase K and incubated at 37°C for 30 minutes to inactivate the alkaline phosphatase. The dephosphorylated DNA was purified by phenol and chloroform/isoamyl alcohol extraction and ethanol precipitation.

2.2.4 DNA ligation

Insert and vector DNA with cohesive ends were ligated using T4 DNA ligase (NBL) in 1X ligation buffer (5 mM Tris-HCl pH 7.8, 1 mM $MgCl_2$, 0.1 mM

dithiothreitol, 0.1 mM ATP and 10 μg/ml BSA). The mixture was incubated at room temperature for 60 minutes and then incubated at 4 °C overnight. The ligated DNA was transformed into *E.coli* the following day or stored at -20 °C until it was required.

2.3 Transformation of *E.coli*

2.3.1 Preparation of competent *E.coli* cells

E.coli strains DE3, TG1, p*Lys*S and DH11S were made competent according to Nishimura *et al.* (1990). 100 ml medium A culture (LB broth supplemented with 10 mM $MgSO_47H_2O$ and 0.2% glucose) was inoculated with 0.5 ml of over-night culture and incubated at 37°C until mid-logarithmic phase ($OD_{600} \sim 0.6$). The cells were kept on ice for 10 minutes, then pelleted at 2000 g for 10 minutes at 4°C. The cells were resuspended gently in 0.5 ml of medium A, pre-cooled on ice. 2.5 ml of storage solution medium B (36% glycerol, 12% PEG-6000 and 12 mM $MgSO_47H_2O$ in L broth) was then added and mixed well without vortexing. The competent cells were aliquotted in Eppendorf tubes and stored at -70°C until use.

2.3.2 Transformation

Competent *E.coli* cell strains (DE3, TG1, p*Lys*S and DH11S) were transformed with plasmid DNA molecules. The frozen competent cells were thawed on ice, and a 100 μl aliquot was added to 10 μl of ligation mixture of plasmid in TE (50-80 ng DNA) and incubated on ice for 30 minutes. The cells were then subjected to a heat-shock at 37°C for 45 seconds, chilled on ice for 2 minutes, supplemented with 400 μl LB and incubated at 37°C for 45 minutes. 50 μl and 200 μl aliquots of the transformed cells were spread onto LB agar plates containing 100 μg/ml ampicillin. Transformation

frequency were calculated on the basis of colony counts after over-night incubation at 37°C.

2.3.3 Preparation of *E.coli* glycerol stocks

Overnight cultures of *E.coli* were mixed with an equal volume of sterile 6:4 LB-medium/glycerol (V/V) and aliquotted into several sterile tubes, some of which were stored at -20 °C for immediate use while others were kept at -70 °C.

2.4 Preparation of plasmid DNA from *E.coli*

2.4.1 Small-scale preparation of double-stranded DNA

Plasmid DNA was prepared from an overnight culture of *E.coli* by the rapid, small-scale, alkaline-lysis method (mini-prep) or sometimes by the QIAprep plasmid preparation system as described in the manufacturer's manual (Qiagen).

3-5 ml of overnight bacterial culture was spun down at 13000 rpm for 30 seconds and the supernatant was removed. The cell pellet was resuspended in 100 µl GTE (50 mM glucose, 10 mM EDTA and 25 mM Tris pH 8.0) by vortexing and was incubated for 5 minutes at room temperature. 200 µl of freshly made 0.2 M NaOH, 1.0% SDS was added, the solution was mixed by inverting the tube. 150 µl of ice-cold 3M K/5M Ac were then added, the suspension was mixed by gently inverting the tube and incubated on ice for 5 minutes. The cell debris were spun down for 10 minutes and the supernatant was transferred to a clean tube. The plasmid DNA in the supernatant was extracted twice with equal volume of phenol, once with chloroform/isoamyl alcohol and was then ethanol precipitated. The DNA was resuspended in 50 µl TE, RNase A was added to a final concentration of 100 µg/ml and incubated for 30 minutes at 37 °C. After the RNase

A treatment, the DNA was phenol and chloroform purified, ethanol precipitated and the pellet was resuspended in the desired volume of TE or H_2O.

2.4.2 Large-scale preparation of double-stranded DNA

50 µl of plasmid glycerol were added to 100 ml LB culture containing 100 µg/ml Ampicillin, and grown at 37 °C overnight. The overnight culture was pelleted by centrifugating at 5000 rpm for 10 minutes at 4 °C. The pelleted cells were washed by resuspending in 50 ml ice-cold GTE, centrifugated at 5000 rpm for 10 minutes at 4 °C, and all residual GTE was removed. Further centrifugation at 3000 rpm for 1 minutes was applied if much supernatant remained. The washed cell pellet was completely resuspended in 9 ml ice-cold GTE. 18 ml of 0.2 M NaOH, 1% SDS was added to the suspension and mixed immediately and thoroughly by inversion. 13.5 ml of 3M K/5M Ac was added and mixed immediately and thoroughly by inversion then shaking. This mixture was incubated on ice for 5 minutes and spun at 6000 rpm for 20 minutes at 4 °C. The supernatant was collected and divided equally into 20 ml Oakridge tubes, extracted with a 50% volume of buffered phenol and chloroform/isoamyl alcohol, and precipitated with an equal volume of propan-2-ol at -20 °C for 60 minutes. The DNA was resuspended in TE and transferred to a microcentrifuge tube, RNase A was added to a final concentration of 100 µg/ml and incubated for 30 minutes at 37 °C. After the RNase A treatment the DNA was phenol and chloroform purified, ethanol precipitated and the pellet was resuspended in total of 1 ml TE. The DNA was then precipitated with 0.6 volume of 20% PEG6000 in 2.5 M NaCl overnight at 4 °C or for 60 minutes on ice. Finally, the DNA was resuspended in an appropriate volume TE and the concentration was determined by A_{260}.

2.5 Amplification of DNA fragments by polymerase chain reaction (PCR)

PCR is the *in vitro* enzymatic amplification of a specific DNA segment by repeated cycles of extension of two oligonucleotide primers defined as the upstream and downstream ends of the desired PCR product. PCR entails mixing template DNA, two appropriate oligonucleotide primers, Taq DNA polymerase, deoxyribonucleoside triphophates (dNTPs), and buffer. Once assembled, the mixture is repeatedly cycled through temperatures that effect denaturation, annealing, and synthesis to amplify, exponentially, a product of specific size and sequence. All PCRs in this study were performed at 1.5mM $MgCl_2$. The annealing temperature was calculated using the following formula:

$$\mathbf{T_{ann} = 61.2 + 0.41(\%GC) - 500/L}$$

Where %GC is the absolute value of the percentage of G+C bases in the oligo and L is the total number of bases in the oligo. For the preparation of GH1 mutants, PCR was carried out in 20 µl final volume. The reaction contained 200 µM dNTPs, 1 µM T7 promoter primer, 1 µM T7 terminator primer, 10 ng template, 0.5% Tween and 1U Taq DNA polymerase enzyme in 1 X digestion buffer containing 1.5mM $MgCl_2$. For the first cycle, template DNA was denatured at 95 °C for 4 minutes, thereafter, 95 °C for 1 minutes. The annealing temperature was 56 °C for 1 minute and extension of the oligomers proceeded at 74 °C for 1 minute. For each PCR, 25 consecutive cycles were performed. The PCR products were assayed by running a sample in an agarose gel.

PCR was also employed as a means to rapidly screen transformants. Here, the template is provided by a colony grown on LB agar plate which is picked with a toothpick into the PCR mix.

2.6 Agarose gel electrophoresis of DNA

DNA restriction digests, amplified DNA prepared by PCR, single-stranded DNA and double-stranded DNA preparations were analyzed by horizontal electrophoresis in a chamber of 8.5 cm X 10 cm dimensions. Nusive agarose (fragment ≤ 1 kb) or SeaKem agarose (fragment ≥ 1 kb) (FMC) was dissolved at the desired percentage in 50 ml of 1 X TBE containing 0.5 μg/ml of EtBr by boiling. The gel was cast with an appropriate comb inserted and allowed to solidify. DNA samples were dissolved in 1 X TBE loading buffer (5 X stock solution: 40% sucrose, 5% glycerol, 4 mM EDTA, 0.05% bromophenol blue). The agarose gel was run in 1 X TBE containing 0.5 μg/ml EtBr at a constant 50 V until the desired resolution was achieved. DNA was detected on an UV transilluminator.

2.7 Determination of DNA concentration

An accurate determination of the DNA concentration was achieved by spectroscopy at 260 nm wavelength taking 1 optical density (OD) unit as equivalent to 50 μg/ml of double-stranded DNA or 20 μg/ml of single-stranded oligonucleotide (Sambrook *et al.*, 1989). If the OD_{260}/OD_{280} was significantly less than 1.8, then it was necessary to further purify the DNA by phenol and chloroform/isoamyl alcohol extraction and ethanol precipitation.

When the amount of DNA available was limiting, its approximate concentration was estimated by dot blot. 2 μl of DNA was spotted onto DE81 chromatography paper (Whatman) along with a dilution series of known concentration standards (from 10 to 100 ng/μl). The paper was then stained with 1X TBE containing 0.5 μg/ml EtBr for 10 minutes, and washed with 1X TBE without EtBr for 10 minutes three times. After

washing, the unknown DNA concentration was estimated from EtBr fluorescence intensity by comparing to the standard DNA series.

Alternatively, the approximate concentration of DNA was estimated by running a sample on an agarose gel together with a dilution series of known concentration DNA. The unknown concentration was estimated from the EtBr fluorescence intensity relative to the intensity of the calibration samples.

2.8 Sephadex G-25/G-50 column chromatography

Sephadex column chromatography was employed to purify DNA from unincorporated ^{32}P after radioactive labelling. A 1 ml syringe was filled in Sephadex G-25/G-50 and was then subject to a 1.6 K spin for 1 minute to pack the column. After packing, TE or H_2O was used to wash the Sephadex in 2 X 1.6 K spins for 1 minute. A third wash was spun at 1.6 K for 6 minutes (5 minutes for G-25). After this wash, the sample was loaded on the top of the column as soon as possible and the column was spun at 1.6 K for 6 minutes. The eluted sample was collected and purified by phenol and chloroform/isoamyl alcohol extraction and ethanol precipitation.

2.9 Preparation of 5′-end labelled primers

2.5 µl of 10 µM T7 promoter primer (25 pmol), 1 µl of 10X labelling buffer (0.5 M TrisHCl pH7.6, 0.1 M $MgCl_2$, 50 mM DTT, 1 mM spermidine, and 1 mM EDTA), 1 µl of 10 U/µl T4 polynucleotide kinase (Amersham), 4 µl [γ-^{32}P]ATP and 1.5 µl H_2O were mixed and incubated at 37 °C for 60 minutes. 115 µl of H_2O were then added and the sample was incubated at 68 °C for 20 minutes to inactivate the T4 polynucleotide

kinase. Labelled primers were purified by chromatography on Sephadex G-25 and the labelled sample was stored at -20 °C.

2.10 Sequencing double-stranded plasmid

DNA sequencing of double-stranded plasmid was performed by using the dideoxy chain-termination method (Sanger *et al.*, 1977).

0.5 pmole of plasmid DNA was denatured in 0.2 M NaOH for 10 minutes at room temperature. The solution was neutralized by adding a 1/5 volume of 5 M ammonium acetate pH7.5, and the DNA was precipitated with 2.5 volumes of ethanol. After washing with 70% ethanol twice, the dry pellet was dissolved in distilled water to a concentration 0.05 pmole/µl and used as soon as possible.

Sequencing reactions were performed in two steps. The first step was to anneal primers to the template: in a 10 µl solution, 1.25 µl of labelled T7 DNA primer (0.25 pmole), 5 µl of single-stranded DNA template (0.25 pmole), 2 µl of 5X sequencing reaction buffer (200 mM Tris-HCl pH7.5, 100 mM $MgCl_2$, and 250 mM NaCl), and 1.75 µl H_2O were mixed together in a PCR tube and incubated at 65 °C for 3 minutes and then cooled down to 30 °C at a rate of 75 seconds/ °C. The next step was extension and termination: 1 µl of 0.1 M DTT, 4.7 µl of 10 mM Tris pH7.5, and 0.3 µl of T7 DNA polymerase (Pharmacia) were added to the annealing mix. The mix was split into four, 3.5 µl aliquots and placed in 4 fresh PCR tubes, each of which contained 2.5 µl of 80 µM dNTPs, 50 mM NaCl and one of the four dideoxy termination nucleotides. These samples were incubated at 37 °C for 5 minutes and then for 15 minutes at 50 °C. The reaction was stopped by adding 1 volume of sequencing stop mix (90% de-ionized Formamide (BDH), 9mM EDTA pH8, 1X TBE, 0.04% Xylene cyanol and 0.04% Bromophenol blue).

2.11 Denaturing acrylamide DNA sequencing gel

DNA sequencing reactions were analyzed on 48 x 20 cm or 50 x 38 cm, 6% acrylamide gels (19:1 acrylamide : N, N′-methylene bisacrylamide) containing 8 M urea. The gel was cast by the addition of 0.04% fresh ammonium persuphate and 0.12% TEMED in 1X TBE, mixed well and poured between the assembled cleaned glass plates. The gel was pre-run for at least 20 minutes before loading the samples. Sequencing reaction samples were heated to 95 °C for 5 minutes before loading and the gel was run in 1X TBE buffer at a constant power setting of 50 W (for 48 x 20 cm) or 75 W (for 50 x 38) for a period of time depending on the size of DNA fragments. At the end of the electrophoresis run, the gel was fixed in 10% glacial acetic acid and 12% methanol for 15 minutes, transferred to 17 mm chromatography paper (Whatman) and dried at 80 °C under vacuum for 1 hour. The dry gel was exposed to X-ray film (Fuji) at -70 °C with intensifying screens.

2.12 Northern blotting and hybridization

2.12.1 Sample treatment and electrophoresis

Northern gels were made by adding agarose to 1X FRB (0.4 M MOPS, 0.1 M NaOAc3H_2O, and 0.02 M EDTA) in DEPC-treated H_2O and boiling. When the solution had cooled to 55 °C, 0.66 M formaldehyde was added and the gel was cast in a tank which had been presoaked in H_2O_2 for 10 minutes and washed with DEPC-treated H_2O. RNA from cell pellets was resuspended in RNA sample buffer (50 mM TrisHCl pH 8, 2 mM EDTA, 1% β-mercaptoethanol, 1% SDS and 10% glycerol). 4.9 µl of purified RNA was added to 10.1 µl sample loading buffer (50% formamide, 1X FRB, 17.8% formaldehyde and 0.01% bromophenol blue). Samples were heated at 66 °C for 15 minutes to denature RNA secondary structure. The samples were loaded onto the

formaldehyde-containing gel and run at 10 V/cm in recirculated formaldehyde running buffer (1X FRB and 18% formaldehyde).

2.12.2 Transfer

The gel was carefully removed from the tank and washed twice with 300-500 ml DEPC-treated H_2O in an RNase-free tray for 5-10 minutes to remove any agarose and excess formaldehyde. The gel was trimmed of its edges and transferred overnight by capillary onto Hybond-N filter in 20X SSC at room temperature. The transfer was assembled by the following steps: (i) place a piece of Whatman 3MM paper on a glass plate to form a support in a box; (ii) place the gel upside down on the Whatman 3MM paper; (iii) surround the gel with Parafilm; (iv) cover the gel with Hybond-N nitrocellulose filter; (v) place a piece of wet 3MM paper; (vi) place the paper towels (5-8 cm high) on top of the paper; (vii) put a glass plate on top of the stack; (viii) finally put an empty 1000 ml Duran bottle on the top. Next day, the filter was then UV-crosslinked for 3-5 minutes.

2.12.3 Hybridization

Both pre-hybridization and hybridization were carried out in a Hybaid oven at 68 ^{o}C. The filter was incubated in 25 ml of pre-hybridization buffer (3X SSC, 10 mM EDTA, 0.1% SDS, 0.2% PVP, 0.2% Ficoll, 0.2% BSA, 0.5 mg/ml heparin and 0.1 mg/ml denatured salmon sperm DNA) in a hybridization bottle (230 X 35 mm) for 2-3 hours. Pre- hybridization solution was drained off and replaced by 25 ml of hybridization buffer (Pre-hybridization buffer + 9% Dextran sulphate). ^{32}P-labelled DNA probe was denatured by boiling and was added to the hybridization buffer and then incubated overnight.

2.12.4 Preparation of ^{32}P-labelled probes

DNA fragments chosen as probes were labelled with [α-^{32}P] dCTP by random priming of DNA hexanucleotides on single-stranded DNA templates by Klenow extension.

Approximately 20 ng of DNA was made up to 29 µl in distilled water and boiled for 3 minutes. The denatured DNA was added to a premix which contained 10 µl OLB buffer, 2 µl BSA (10 mg/ml), 1µl each of 10 mM dATP, dTTP and dGTP, 5 µl of [α-^{32}P] dCTP (10 µCi/µl), and 1 unit of Klenow. The OLB buffer was made from the following components: 1 volume of Solution A (1 ml Solution O, 18 µl β-mercaptoethanol), 2.5 volumes of 2 M Hepes pH 6.6 and 1.5 volumes of random Hexadeoxyribonucleotides @ 4.5 mg/ml (Pharmacia). Solution O was made from 1.25 M Tris-Cl (pH7.5) and 0.125 M $MgCl_2$.

The reaction mix was incubated for 30 minutes at 19 °C, and then for 30 minutes at 30 °C to allow Klenow extension of the annealed primers. At the end of the reaction, the labelled DNA was purified by G-50 Sephadex chromatography.

2.12.5 Filter Washing

Overnight-hybridized filters were washed of excess probe and non-specific hybrids by 4 X 20 minutes washes with 2 X SSC, 0.1% SDS, followed by a further 2 X 20 minutes washes with 0.1 X SSC, 0.1% SDS. Finally, the filter was blotted dry with paper towels before sealing into a thin plastic bag for autoradiography.

2.13 Native DNA acrylamide gels

Native DNA acrylamide gels were needed to assess the digested products of chromatosome protection analysis. A 15 cm X 20 cm acrylamide gel containing 7% acrylamide (19:1 acrylamide : N,N' methylene bisacrylamide), 1 X TBE, 0.07% ammonium persulphate and 0.035% TEMED was prepared at least 2 to 3 hours prior to use. Micrococcal nuclease-digested chromatosomal DNA, resuspended in 10 to 20 µl of TBE loading buffer (see Chapter 2.6), was loaded on to the gel and run at a constant 100 V for 2 hours in 1 X TBE running buffer. The gel was stained with 1X TBE containing 0.5 µg/ml of EtBr for 20 minutes and the DNA was detected by UV transillumination.

2.14 Preparation of the recombinant globular domains of linker histones H1 or H5

2.14.1 Cell growth and acid extraction

The recombinant globular domain of linker histones H1 (GH1) or H5 (GH5) were prepared by expression in *E.coli*.

160 µl of GH1 or GH5 overnight LB culture was inoculated into 160 ml LB containing 100 µg/ml ampicillin. When the OD_{600} reached 1.0, 1 mM IPTG was added to induce GH1 or GH5 expression. After 3 hours, the culture was collected in a 250 ml bottle and left on ice for 10 minutes. The pellet was collected by a 5K spin for 10 minutes at 4 °C. The pellet was resuspended in 25 ml grinding buffer with homogenization and sonication. The suspension was spun 16K for 15 minutes with at 4°C. The supernatant was collected and 70% perchloric acid (PCA) was added to a final concentration of 5% (V/V). The PCA extraction was allowed to progress on ice for 1 hour whereupon the insoluble material was pelleted by spinning 16 K for 30 minutes (JA20) at 4°C. The supernatants were pooled and the protein precipitated by the addition

of 6 volumes of acetone and incubation at –20°C overnight. The precipitate was collected by centrifugation and the pellets were washed 3 times in acetone and dried in vacuum.

2.14.2 Fractionation of acid-soluble proteins

Fractionation of 5% PCA-soluble proteins was achieved by ion-exchange chromatography using the cation exchanger sephadex CMC-25 (Pharmacia). The dried pellet of acid-soluble proteins was dissolved in 1.2 ml of H_2O containing 0.1 mM PMSF. 3 ml of CMC-25 was packed into a column and equilibrated with 10 ml of 20 mM phosphate. 1 ml of sample, containing 490 µl acid-soluble proteins, 10 µl of 1 M phosphate pH 6.8, and 500 µl of 20 mM phosphate was loaded into the top of the column. The column was then washed with (i) 5 ml of 20 mM and (ii) 5 ml of 20 mM phosphate, 200 mM NaCl, (iii) 5 ml of 20 mM phosphate, 700 mM NaCl. Fractions were collected and the absorbance measured at 280 nm. GH1 and GH5 eluted in the 700 mM NaCl wash. Peak fractions were pooled and stored at –20°C.

2.15 Gel electrophoresis of proteins

Proteins were routinely assayed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) in a Tris-glycine discontinuous buffer system according to the method described in Sambrook *et al.* (1989).

2.15.1 Preparation of SDS-polyacrylamide gels and protein samples

A 9 X 9 cm slab gel of 1 mm thickness was routinely prepared. 10 ml of 15% separating gel was made by mixing 5 ml 30% acrylamide (29:1, acrylamide : N,N'-methylenebisacrylamide), 2.5 ml 1.5 M TrisHCl (pH 8.8), 100 µl 10% SDS, 100 µl 10%

ammonium persulfate (freshly prepared), 4 µl N,N,N',N'-tetramethyl-ethylenediamine (TEMED), and 2.5 ml H_2O. After mixing, 6.85 ml of the mixture was poured to the space between the glass plates, overlaid with butanol, saturated with 0.375 M Tris-HCl pH 8.8, and allowed to polymerize for at least 1 hour. Once the separating gel was set, the saturated butanol was removed and a 5% stacking gel was poured on top of the separating gel. The 5% stacking gel contained 5% acrylamide, 130 mM Tris-HCl pH 6.8, 0.1% SDS, 0.01% ammonium persulfate, and 0.1% TEMED. A gel comb was inserted into the top of the stacking gel. After polymerization, the comb was carefully removed and the gel was mounted into the electrophoresis apparatus. The top and bottom reservoirs were then filled with Tris-glycine electrophoresis buffer (25 mM Tris, 250 mM glycine, and 0.1% SDS).

Purified recombinant protein or total cell protein were resuspended in loading buffer (62.5 mM Tris-HCl pH 6.8, 5% glycerol, 0.005% bromophenol blue, 1 % SDS, 0.5% β-mercaptoethanol). The resuspended proteins were heated at 95°C for 5 minutes before they were loaded onto the gel. Electrophoresis was at a constant 100V. When the tracking dye had reached the separating gel, the voltage was increased to 250V.

2.15.2 Detection of proteins in SDS-polyacrylamide gel

After electrophoresis, the gel was immersed in Coomassie blue staining solution (0.05% Coomassie brilliant blue R250, 45% methanol, 9% acetic acid) for 1 to 2 hour(s) at room temperature. Excess stain was removed from the gel by transferring it into a destaining solution (10% methanol, 10% acetic acid) for up to 14 hours in the presence of a small bundle of raw silk. After destaining, the gel was stored in destaining buffer in a sealed plastic bag or dried onto a piece of Whatman 3 MM paper.

2.16 Determination of protein concentration

GH1 and GH5 protein concentrations were measured by UV spectroscopy on a spectrophotometer calibrated with K_2CrO_4 in 0.5 N KOH, using semi-micro quartz cells of 10 mm path-length. An absorption value for a 1 mg/ml solution of GH5 in 1 mM phosphate pH 7.4 at 278 nm wavelength is 0.49, and 0.117 for a 1 mg/ml solution of GH1.

2.17 Preparation of long-range sequencing ladders

Long C/T sequencing standards were prepared as size markers for monomer extension gels.

In a PCR tube, 1 pmole of M13-20 primer (not labelled) was mixed with 1 pmole of M13mp18 single-stranded DNA (USB) in 1 X sequencing buffer (40 mM Tris pH 7.5, 20 mM $MgCl_2$, and 50 mM NaCl) and heated to 65°C for 4 minutes then ramped to 30°C at 75 sec/°C. After this annealling process, 4 µl of 10 mM Tris pH7.5, 2 µl of 0.1 M DTT, 2 µl of X5 labelling dNTP mix (7.5 µM dCTP, dGTP, and dTTP), 3 µl of α^{35}SdATP (600 Ci/mmol), and 1 µl of T7 DNA polymerase (8 units/µl) were added to the annealling mixture and incubated at room temperature for 10 minutes.

The labelling mix was split 2 X 14 µl into 2 PCR tubes which contained 10 µl of C term or T term (50 mM NaCl, 80 µM dNTPs, 8 µM ddC or ddT), respectively, and the extension was proceeded at 37°C for 5 minutes, 50°C for 15 minutes, and then 30°C for 10 minutes. In the end of the reaction, the mixtures were phenol and chloroform purified, ethanol precipitated, and the pellet was resuspended in 40 µl sequence stop mix (SSM, 9 mM EDTA pH 8, 1 X TBE, 0.04% Xylene Cyanol, 0.04% bromophenol blue, and 90% deionized formamide (BDH)).

2.18 Preparation of ^{32}P end-labelled DNA size markers

Phage lambda DNA was cleaved by HinFI or DdeI restriction enzymes, purified by ethanol precipitation and then 5′ end labelled with T4 polynuclease kinase to be used as size markers for monomer extension gels.

300 ng of HinFI or DdeI-digested DNA was mixed with 1 X 5′ kinase buffer (10 X buffer is 500mM Tris pH 7.6, 100 mM $MgCl_2$, 50 mM DTT, 1mM spermidine, and 1 mM EDTA), 3 µl of γ^{32}P-ATP, and 10 units of T4 polynuclease kinase, and incubated at 37°C for 60 minutes. In the end of the reaction, the mixtures were phenol, chloroform extracted, followed by Sephadex G-50 column purification and ethanol precipitation. Finally, the pellet was resuspended in 40 µl SSM and stored at –20°C.

Chapter 3 Expression of Globular Domain of the Chicken Linker Histone H1 in *E.coli*

3.1 Introduction

Protein translation is a complex and free-energy-demanding process comprising three stages: initiation, elongation and termination. The primary rate-limiting step controlling the translation of mRNA in *E.coli* is the process of initiation (Figure 3-1)(Gold, 1988; McCarthy & Gualerzi, 1990; Gualerzi & Pon, 1990; Voorma, 1996; Makrides, 1996). The first steps in this procedure involve the interaction of mRNA and initiator tRNA (fMet-tRNA^{fMet}) with the 30S ribosomal subunit, an assembly of 21 proteins, 16S rRNA, initiation factors (IF1, IF2 and IF3) and a molecule of GTP. The resulting 30S complex is subsequently transformed into a 70S initiation complex by association with a 50S ribosomal unit, an assembly of 31 proteins and 23S rRNA. As a result of these interactions, the fMet-tRNA^{fMet} becomes correctly located on the ribosomal P-site and the junction with the 50S ribosomal subunit then takes place, triggering GTP hydrolysis and release of the initiation factors and GDP. The 70S initiation complex can then interact with the incoming aminoacyl-tRNA and begin the elongation stage of translation (Figure 3-1).

3.1.1 The role of initiation factors

Prokaryotic initiation factors stimulate the rate of 30S initiation complex formation. The main function of initiation factors is to affect kinetically the formation and dissociation of the codon-anticodon interaction at the P-site of the 30S subunit and to influence 30S initiation complexes entering the elongation cycle after association with the 50S subunit. From fluorescence polarisation and centrifugation studies, it has been shown that the 30S subunit has a single high-affinity site for each factor (Gualerzi and Pon, 1990; Voorma, 1996). IF2 binds fMet-tRNA^{fMet} and GTP and positions fMet-tRNA^{fMet} into the P-site of the 30S subunit. It favours the binding of aminoacyl-

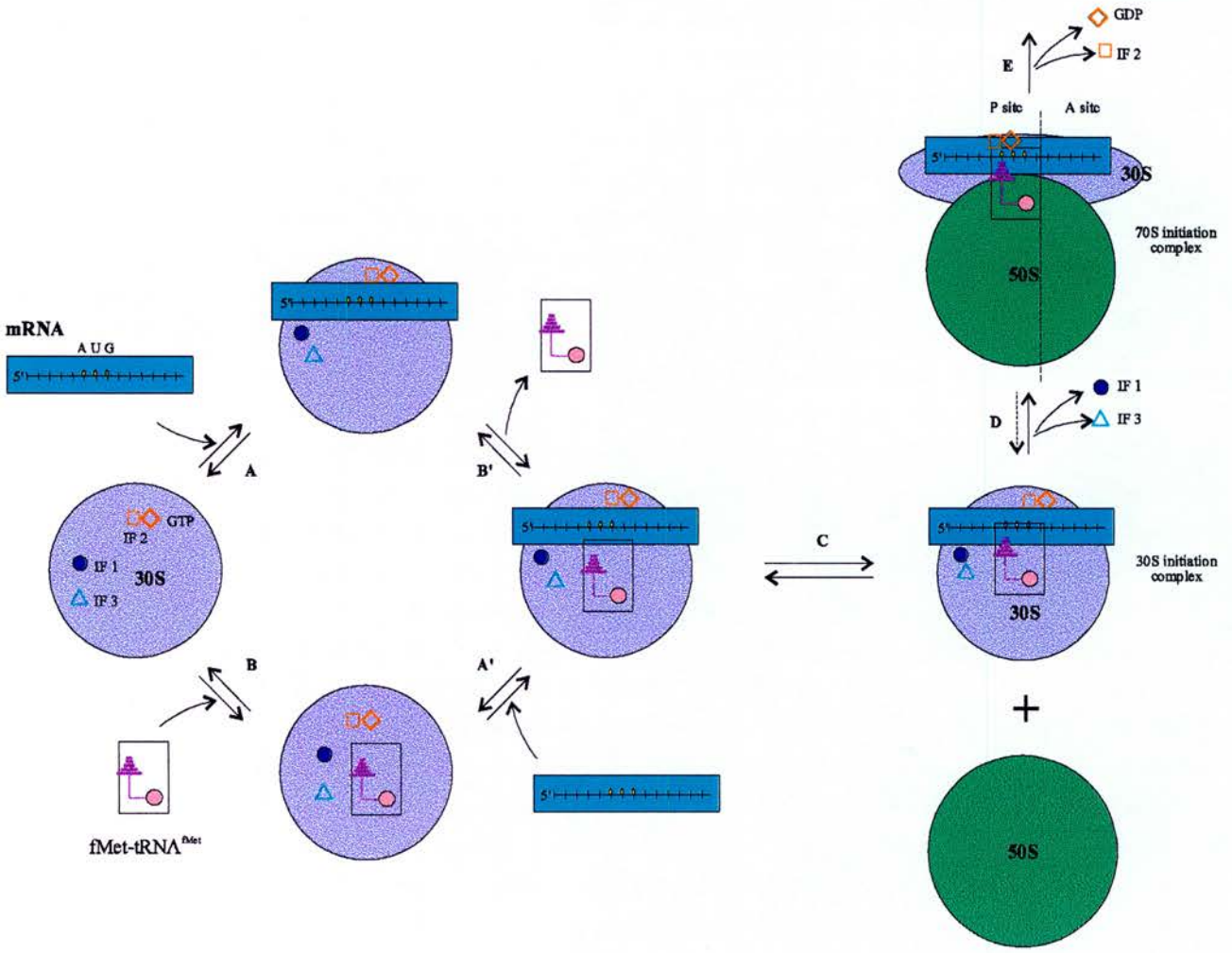


Figure 3-1 Scheme of translation initiation (adapted from Gualerzi & Pon, 1990; McCarthy & Gualerzi, 1990). Step A, A', B, and B' are in rapid equilibrium. Step C represents the first-order, rate-limiting rearrangement of the ternary preinitiation complex kinetically controlled in both directions by the IFs. Step D represents the virtually irreversible subunit association giving rise to the 70S initiation complex.

tRNAs with blocked α–amino groups and thereby helps to exclude noninitiator tRNAs from binding to the 30S subunit. It is a kinetic effector of 30S and 70S initiation complex formation. IF2 can also function as a ribosome-dependent GTPase which serves to more rapidly eject the factor following 70S initiation complex formation. IF3 binds to 30S subunits, acts as an anti-association factor, and increases the affinity of 30S subunits for IF1 and IF2. It inspects the correctness of the fMet-tRNA^{fMet} anticodon stem and the P-site codon-anticodon interaction and then effects ejection of noninitiator tRNAs or interactions with noninitiation codons (Hartz *et al.*, 1989). In considering the sequence of events in the formation of the 30S initiation complex, Gualerzi *et al.* (1986) have argued that the complex could be assembled by first binding either the mRNA or the fMet-tRNA. A weak ternary complex, fMet-tRNA•30S•mRNA, the rapid formation of which requires the presence of IF3, would be transformed into a stable complex upon interaction of the initiator tRNA with its cognate codon, AUG. Therefore, the IF3 on the one hand accelerates the formation of the stable initiation complex and on the other hand discriminates against artificial complexes, such as between N-acetyl-Phe-tRNA and poly(U)-programmed 30S. IF2 and IF3 accelerate the locking and unlocking of the codon-anticodon interaction at the P-site while favouring the formation of the correct over the incorrect and the dissociation of the incorrect over the correct 30S complexes. The function of IF1 remains obscure, but it appears to stimulate both IF2 and IF3 activities. It binds to 30S subunits and is ejected upon subunit association,

In consideration of both ribosome-mRNA interaction and initiation factor function, none of the IFs influence the Shine-Dalgarno (SD) interaction and the affinity of the 30S subunit for mRNAs with or without SD sequence, because no effect on the association constants of the mRNA-30S complexes has been detected (Calogero *et al.*, 1988; Canonaco *et al.*, 1989; Laughrea & Tam, 1990). Although the IFs have no detectable effect on the constants of the binary complexes between 30S subunits and mRNAs, their presence may well affect the position of the mRNA in its ribosomal binding site. Therefore, Canonaco *et al.* (1989) proposed that (i) in the absence of IFs, the mRNA preferentially occupies a ribosomal "standby site" corresponding to the region where the SD interaction takes place and (ii) in the presence of the IFs, the

mRNA is shifted toward another ribosomal site, possibly closer to the ribosomal P-site where the IFs exert their kinetic influence on codon-anticodon interaction.

3.1.2 The features of initiator tRNA

Escherichia coli contains a main and a minor form of initiator tRNA, $tRNA^{Metf1}$ (~75%) and $tRNA^{Metf2}$ (~25%). Both tRNAs contain 77 nucleotides and differ in the presence of either 7-methyl-G or A at position 47. The initiator tRNA has a similar crystal structure and the same CAU anticodon as the $tRNA^{Met}$ used in elongation and is aminoacylated by the same synthetase, which recognises primarily the bases of the anticodon (Pelka & Schulman, 1986). However, the unique features of tRNA are (i) the presence of three consecutive GC base pairs in the anticodon stem that confer rigidity and regularity to the helix and a particular conformation to the anticodon loop, which targets the initiator tRNA to the ribosomal P-site; (ii) the recognition by N^{10}-formyltetrahydrofolate-Met-tRNA transformylase which modifies the α-NH_2 group, that results in a 6-10-fold stimulation of protein synthesis in *E.coli* extracts; (iii) the absence of a Watson-Crick base pair at the end of the amino acid acceptor stem which is responsible for the resistance of fMet-tRNA to the action of peptidyl-tRNA hydrolase that leads to weak interaction with EFTu-GTP (Gualerzi & Pon, 1990).

Under physiological conditions, the conversion of the 30S-mRNA binary complex to a stable 70S ternary complex is accomplished by binding of initiator tRNA and initiation factors. After the association of 50S and 30S unit, the fMet-tRNA is located in the ribosomal P-site, where it can react with the incoming aminoacyl-tRNA encoded by the second mRNA codon in the ribosomal A-site. The formation of the first peptide bond marks the transition to the elongation stage of translation (McCarthy & Gualerzi, 1990).

3.1.3 Structural features of mRNA

Of all the components involved in the initiation of translation, the sequence of the mRNA constitutes the primary variable and as such has the potential to exert a

major influence upon the efficiency of protein production. In this context, the most important parts of the mRNA are the sequences which comprise the translational initiation region (TIR). Initiation is dictated by the start codon (most commonly AUG (90%), less frequently GUG (8%) or UUG (1%), and in one case AUU) which in natural mRNAs is flanked on its 5′ side by a Shine-Dalgarno sequence, a region intimately involved in directing initiation by virtue of its capacity to base pair with the 3′ end of the 16S RNA component of the 30S ribosomal unit (the anti-Shine-Dalgarno sequence). The area from the 5′ Shine-Dalgarno sequence to the downstream start codon is believed to have great impact on translational efficiency due to its unique structural features (Scherer *et al.*, 1980; Gold *et al.*, 1981; Hui *et al.*, 1984; Schneider *et al.*, 1986).

3.1.3.1 Shine-Dalgarno region

Shine & Dalgarno (1974) were the first to identify a sequence in the ribosome-binding site (RBS) of bacterial phage mRNAs containing all or a substantial part of the sequence 5′-GGAGGU-3′ subsequently called the Shine-Dalgarno (SD) site. They proposed that this purine-rich sequence in the mRNA interacts with a complementary sequence at the 3′ end of the 16S rRNA, now called the anti-Shine-Dalgarno (ASD) sequence, thereby promoting the binding of mRNA to the 30S ribosomal subunit. The hypothesis was confirmed by the isolation from initiation complexes of a nuclease-resistant dimer between the RBS and ASD (Steitz & Jakes, 1975). The importance of the SD interaction with the 16S rRNA during the initiation of translation is now well established (Hui & de Boer, 1987; Jacob *et al.*, 1987; McCarthy & Brimacombe, 1994). The spacing between the SD site and the initiating AUG codon can vary from 5 to 13 nucleotides, and it influences the efficiency of translational initiation (Gold, 1988). Extensive studies have been carried out to determine the optimal nucleotide sequence of the Shine-Dalgarno region and the most effective spacing between the SD region and start codon which is 7-9 nucleotides (Gold, 1988; Ringquist *et al.*, 1992; Barrick *et al.*, 1994; de Smit & van Duin, 1994a; Wilson *et al.*, 1994). The dependence on the length or sequence of the Shine-Dalgarno sequence may suggest that there is a precise physical relationship between the 3′ end of 16S rRNA and the anticodon of the fMet-tRNA^{fMet} bound to the ribosomal P-site (Ringquist *et al.*, 1992;

Chen *et al.*, 1994). Furthermore, from experiments of nucleotide substitutions falling within the region between the SD and start codon, it has been shown that there is a tendency for avoiding nucleotides C and G in this domain on the template, their introduction generally leading to impaired translation (Singer *et al.*, 1981; Gren, 1984). In addition, the nature of the Shine-Dalgarno sequence may also influence the translation efficiency via mRNA stability which affects the interaction between the 30S ribosomal subunit and mRNA (Wagner *et al.*, 1994).

3.1.3.2 Secondary structure of mRNA

Apart from the three classical RBS signals (start codon, SD region and spacing), the structure of the entire translation initiation region may have a substantial effect on translation. Intramolecular base-pairing is often a major factor controlling the initiation efficiency by forming mRNA structures that either hide or expose the recognition elements (Gold, 1988). It is believed that the occlusion of the SD region and/or the AUG codon by a stem-loop structure prevents accessibility to the 30S ribosomal subunits and inhibits translation (Gheysen *et al.*, 1982; Ramesh *et al.*, 1994). Work on the *lamB* gene of *E.coli* first showed that the formation of a base-paired hairpin structure in the *lamB* transcript, which makes the Shine-Dalgarno sequence inaccessible to ribosomes, decreases the translation efficiency (Hall *et al.*, 1982). Many analyses which introduced point mutations in different genes also showed the same result (Tessier *et al.*, 1984; Buell *et al.*, 1985; Freier *et al.*, 1986).

Translational efficiency is strictly correlated with the fraction of mRNA molecules in which the ribosome binding site is unfolded, indicating that initiation is completely dependent on spontaneous unfolding of the entire initiation region (de Smit & van Duin, 1990a). The efficiency of translation is determined by the overall stability of the structure at the RBS, whether the initiation codon itself is base-paired or not. The stability of an RNA secondary structure is usually expressed as its free energy of formation (ΔG^{o}). Since the formation and disruption of helices are in a dynamic equilibrium, the ΔG^{o} is a measure of the fraction of molecules present in the unfold state. Structures weaker than –6 kcal/mol usually do not reduce translational efficiency. Below this threshold, all systems show a ten-fold decrease in expression for

every -1.4 kcal/mol. More recent evidence has shown that highly unstructured 5′-mRNA regions were found in all efficiently expressed genes without any obvious sequence homologies (Helke *et al.*, 1993; Groeneveld *et al.*, 1995). However, by using site-directed mutagenesis, de Smit & van Duin (1994a) varied the extent of the SD-ASD complementarity in the coat-protein gene of bacteriophage MS2 and found that mutations reducing the complementarity by one or two nucleotides diminish translational efficiency only if ribosome binding is impaired by the structure of the messenger. Therefore, in the absence of an inhibitory structure, these mutations have no effect which means a strong SD interaction can provide the ribosome with extra affinity for the messenger, and thus compensate for the inhibition by the secondary structure.

The inhibitory effect of secondary structure on translational initiation is reflected in the fact that natural ribosome binding sites generally display a lower potential than other parts of the messenger to form stable helices (Movva *et al.*, 1980; Cone & Steege, 1985b). Statistical analysis has shown that the absence of stable helices in natural ribosome binding sites is usual (Scherer *et al.*, 1980; Ganoza *et al.*, 1987). The strong preference for A residues, rather than G or C residues, around the initiation region only allows the formation of relatively weak structures. However, the increased base-pairing potential around internal AUGs could not be attributed entirely to the base composition. Some evidence has shown that the function of encoded hairpin structures may relate to the shielding of internal pseudo-initiation sites (Cone & Steege, 1985a; Cone & Steege, 1985b; Ganoza *et al.*, 1987).

3.1.3.3 Further interactions between mRNA and 16S rRNA

The degree of complementarity of the Shine-Dalgarno sequence with the anti-Shine-Dalgarno sequence at the 3′ end of the 16S rRNA, the initiation codon, the spacing between the SD and the start codon as well as secondary structures sequestering the SD or start codon have been shown to be major determinants affecting the translational yield of an individual cistron. However, the non-random distribution of nucleotides located upstream of the SD region and downstream of the initiation codon (Storma *et al.*, 1982) indicates that other primary mRNA sequences might also

be involved directly in the recognition process. Numerous studies have demonstrated that sequence differences in or immediately adjacent to the translational initiation region (TIR) of an mRNA can have profound effects upon translational efficiency.

Statistical analysis and experiments have shown that the translation initiation domains are non-random from –20 to +15 and suggested a role for additional ribosome recognition signals located in this area (Schneider *et al.*, 1986; Dreyfus, 1988). Those nucleotides are also protected against nucleases in initiation complexes (Steitz, 1975). A translational enhancer element has been identified to be located upstream of the SD sequence (McCarthy *et al.*, 1985; Bingham & Busby, 1987; Zhang & Deutscher, 1989; Hartz *et al.*, 1991; Zhang & Deutscher, 1992). In combination with a functional SD region, this element has been used successfully to promote efficient translation of various prokaryotic and eukaryotic genes in bacteria (McCarthy & Gualerzi, 1990; McCarthy & Brimacombe, 1994). A RBS derived from gene 10 of bacteriophage T7 (*g10-L*) caused a pronounced stimulation of expression when placed upstream or downstream of the initiator codon (Olins & Rangwala, 1989). This may suggest that although the *E.coli* TIR contains the essential elements for directing initiation, other mRNA sequences located either 5′ or 3′ to this region, can often have a substantial impact on the efficiency of the process.

It has been suggested that the capacity of these additional sequences to form secondary mRNA structures, or to interact with other regions of the 16S rRNA molecule, may enhance the initiation process, although we have little understanding of the mechanism involved. Sprengart *et al.* (1990) isolated the bacteriophage T7 gene 0.3 on DNA fragments of differing length and cloned them upstream of the mouse dihydrofolate reductase gene in an expression vector to characterise their influence upon the control of translation for this gene's sequence. They found that the TIR's efficiency was highly dependent on nucleotides +15 to +26 downstream of the AUG. This sequence is complementary to nucleotides 1471-1482 of the 16S rRNA (see Figure 3-22), suggesting a second mRNA-rRNA base-pairing contact besides the SD region. Subsequently, mutations in the initial 5′ coding region that reduced or enhanced complementarity between the putative downstream box (DB) in the mRNA

and the anti-downstream box (ADB) in the 16S rRNA, decreased and increased translation, respectively (Faxen *et al.*, 1991; Nagai *et al.*, 1991; Ito *et al.*, 1993).

In a related experiment, Sprengart *et al.* (1996), by introducing nucleotide substitutions into the DB (+9 to +21) and into the SD region of the bacteriophage T7 gene10, found that the DB was not functional when shifted upstream of the initiation codon to the position of a SD region. Furthermore, they observed that nucleotides 1469-1483 of 16S rRNA ('ADB') are complementary to the DB. Optimising this complementarity strongly enhanced translation in the presence or in the absence of a SD region. Therefore, they proposed that the stimulatory interaction between the DB and the ADB places the start codon in close contact with the decoding region of 16S rRNA, thereby mediating independent and efficient initiation of translation. These findings demonstrate convincingly that, in addition to the SD site and the start codon, other sequences in the mRNA can be important for efficient translation.

Due to the fact that the 16S rRNA ADB is located in the penultimate stem of the 16S rRNA (Figure 3-22), which is proposed to be in double-strand coil due to base-pairing to nucleotides 1421-1427, it has been questioned whether the ADB is in fact available for interactions with the mRNA. Studies into the effect of ribosomal protein S2-deficient mutants on the expression rate of the leaderless λ*cI* mRNA indicate that in S2-deficient ribosomes, the 16S rRNA sequence spanning nucleotides 1421-1427 may undergo an alternative intra-molecular base-pairing with nucleotides 1107-1113. This would free the ADB for interactions with the DB on the mRNA (Shean & Gottesman, 1992). However, using hydroxyl radical footprinting it has been shown that the alternative base-pairing between nucleotides 1421-1427 and 1107-1113 is unlikely to occur (Powers & Noller, 1995). Moreover, kinetic toeprinting experiments and *in vivo* expression studies have shown that the DB is dispensable for the initial interaction between the ribosome and the mRNA in the pathway towards formation of a productive initiation complex (Resch *et al.*, 1996). Clearly, the molecular basis for this effect remains to be elucidated.

3.1.4 Aims of this chapter

Overexpression of the globular domain of the chicken linker histone H1 (GH1) in *E.coli* has the potential to allow studies relevant to chromatin structure. For example, this would help me to easily produce large amounts of GH1 without trypsinisation for studies of binding to nucleosome. Furthermore, the pure product is potentially suitable for crystallisation and could allow isotopic labelling for structural determination by neutron scattering. The many advantages of *E.coli* have ensured that it remains a valuable organism for the high level production of recombinant proteins (Nicaud *et al.*, 1986; Marino, 1989; Gold, 1990; Olins & Lee, 1993; Shatzman, 1995; Georgiou, 1996). However, in spite of the extensive knowledge on the genetics and molecular biology of *E.coli*, not every foreign gene can be expressed efficiently in this organism. This may be due to the unique and subtle structural features of the gene sequence, the stability and translational efficiency of mRNA, codon usage, degradation of the protein by the host cell or the potential toxicity of the protein to the host.

In attempting to express the globular domain of chicken linker histone H1 (GH1), it was found that a construct containing the wild-type GH1 coding sequence was very poorly translated in *E.coli*. This was not perhaps unexpected, because the coding sequence for GH1 is particularly GC rich at its 5′ end, a condition which is often detrimental for protein expression in bacteria. By mutating the third position in each of the first 4 codons in the GH1 coding region, with the aim of increasing AT content and optimising codon usage, we attained an elevation of expression by 300-fold.

The aim of this study was to gain an insight into why the sequence changes had such a large effect upon GH1 expression. To this end, I generated a set of 15 mutants comprising all possible combinations of the wild-type and altered codons and compared their capacity to express GH1 with the wild type construct.

3.2 Materials and methods

3.2.1 Plasmid constructs

The coding sequence for the globular domain of the chicken linker histone H1 was prepared from a pUC19 derivative of pCH1aB1BK1 (Sugarman *et al.*, 1983) by PCR. The primers employed possessed non-pairing 5′ termini containing unique restriction sites (BamHI and NdeI); the reverse primer also contained 3 consecutive in-frame termination codons (Figure 3-2A). The purified amplification product was restricted with BamHI and NdeI and cloned into the T7 polymerase-based expression vector pT79 (Figure 3-2B), a derivative of pMW172 (Way *et al.*, 1990). The resulting construct, GH1-WT, contained an in-frame H1 fragment encoding the 76 amino acid GH1 peptide (residues 35-110 of the parent H1) with an additional N-terminal methionine residue (Figure 3-2C).

3.2.2 Site-directed mutagenesis

Site-directed mutagenesis was carried out by PCR of GH1-WT using a degenerate forward primer (Figure 3-3A), designed to introduce the desired sequence changes in random combinations at the 5′ end of the GH1 coding sequence; the reverse primer (Figure 3-3A), was complementary to the T7 termination site in the pT79 expression vector (Figure 3-2B). The PCR product population was digested with BamHI and NdeI, ligated into BamHI/NdeI-cut pT79 and transformed into TG1 host cells (Figure 3-3B). DNAs from numerous individual colonies were sequenced by the dideoxy chain-termination method (Sanger *et al.*, 1977) in order to isolate all 15 GH1 coding sequence variants.

(A)

Forward primer
5'-GG**CATAT**GCCCGCGGGCCCCAGC-3'
NdeI

Reverse primer
5'-*TGATGATGA***GGATCC**TCATCA-3'
stop codons BamHI

(B)

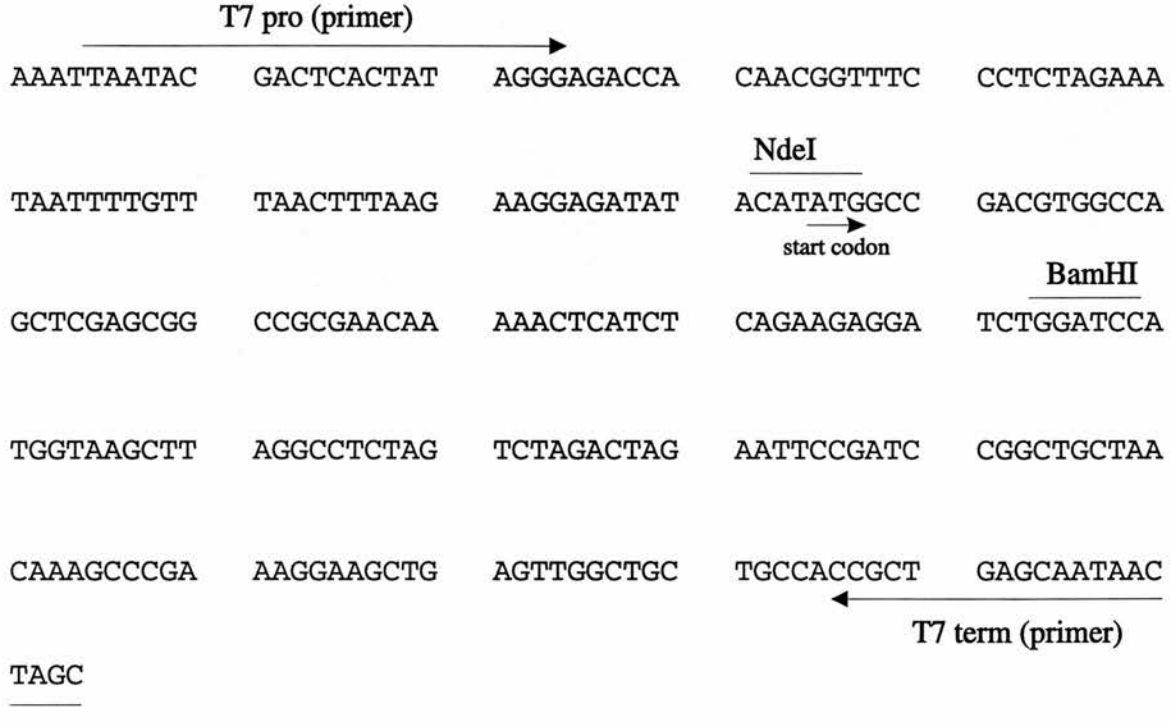


Figure 3-2 (A) The primers employed in the construction of pT79 GH1 plasmid. The stop codons, BamHI and NdeI sites are indicated. (B) The sequence around the cloning region of the expression vector pT79. (C)The coding sequence of globular domain of chicken linker histone H1. The first four codons are indicated. (Adapted from Sugarman *et al.* 1983).

(C)

```
                        1   2   3   4
aag ccc cgc aag CCC GCG GGC CCC AGC GTC ACC GAG CTG ATC ACC AAG GCC
                Pro Ala Gly Pro Ser Val Thr Glu Leu Ile Thr Lys Ala

GTG TCC GCC TCC AAG GAG CGC AAG GGG CTC TCC CTC GCC GCG CTC AAG AAG
Val Ser Ala Ser Lys Glu Arg Lys Gly Leu Ser Leu Ala Ala Leu Lys Lys

GCG CTT GCC GCC CGC GGC TAC GAC GTG GAG AAG AAC AAC AGC CGC ATC AAG
Ala Leu Ala Ala Gly Gly Tyr Asp Val Glu Lys Asn Asn Ser Arg Ile Lys

CTG GGG CTC AAG AGC CTC GTC AGC AAG GGC ACC CTG GTG CAG ACC AAG GGC
Leu Gly Leu Lys Ser Leu Val Ser Lys Gly Thr Leu Val Gln Thr Lys Gly

ACC GGC GCC TCG GGC TCT TTC AAG CTG AAT AAA AAG ccg ggt gag aga
Thr Gly Ala Ser Gly Ser Phe Lys Leu Asn Lys Lys
```

Figure 3-2 (cont.)

(A)

1 2 3 4

Forward primer : degenerate GH1 5'-ATACATATGCC**M**GC**R**GG**Y**CC**M**AGCGTCACCG-3'

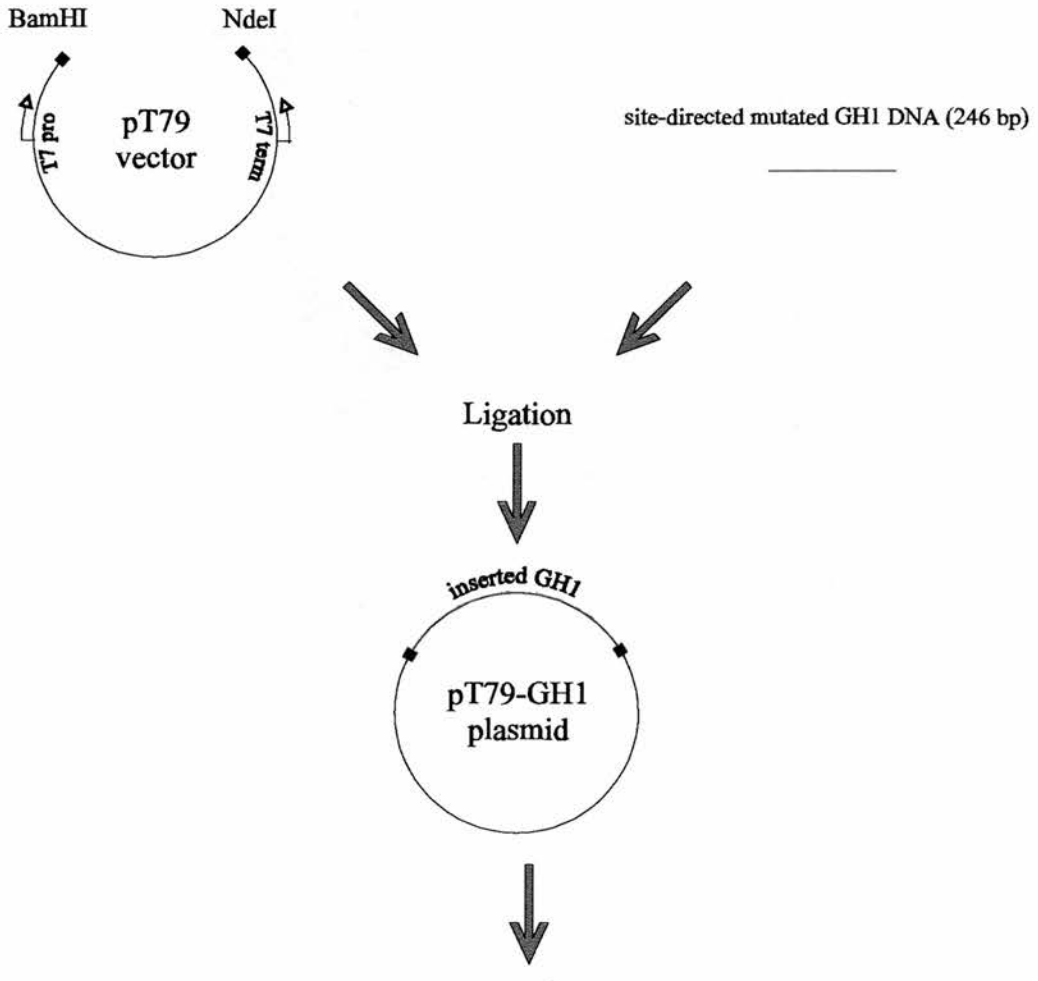
Reverse primer : T7 term 3'-GGCGACTCGTTATTG-5'

M=C or A

R=G or A

Y=C or T

(B)



Transform ligated plasmid into *E.coli* TG1 and screen for transformants

Figure 3-3 (A) Primers for site-directed mutagenesis. (B) The process of ligation and transformation.

3.2.3 GH1 expression

For expression, each of the GH1 plasmid DNA constructs were transformed into DE3 (BL-21) cells. Single colonies were selected and then grown in L broth containing 50 μg/ml ampicillin. Overnight cultures were diluted 1:50 into fresh LB. When the OD_{600} of these cultures reached 1.0, 0.75 mM IPTG was added to induce the expression of GH1. After a further 3 hours, the OD_{600} was again measured and then cells from 1 ml of culture were collected by centrifugation and lysed by resuspending in SDS-gel loading buffer (62.5 mM Tris pH 6.8, 0.5% β-mercaptoethanol, 1% SDS, 5% glycerol and 0.005% bromophenol blue). Total protein from equivalent amounts (OD_{600}) of culture were analysed in 15% SDS-polyacrylamide gels. Coomassie blue stained gels were scanned on a Bio-Rad Imaging densitometer and the data analysed using the Bio-Rad Molecular Analyst software package. GH1 expression levels were normalised by reference to a number of *E.coli* host protein bands.

3.2.4 Northern blotting

Cells collected from 50 μl of induced culture were resuspended in gel loading buffer (20 mM MOPS, 5 mM NaOAc, 1 mM EDTA, 50% formamide, 0.018% formaldehyde, 0.01% bromophenol blue). The RNA in an equivalent amount of each extract, based on the OD_{600} measured at the end of the induction period, was fractionated in a 1% agarose gel containing formaldehyde (Sambrook *et al.*, 1989) (see Chapter 2 for details). After transfer to Hybond-N nylon filter (Amersham), GH1 mRNA was detected by hybridisation using a ^{32}P-labelled DNA probe, specific for the GH1 coding sequence. This DNA probe is generated from the product of restriction enzyme (both NdeI and BamHI) digestion of GH1-WT plasmid.

3.2.5 RNA secondary structure and prediction of free energy of intermolecular interaction

The secondary structures of the 16 GH1 variants and their minimum free energies, and the structures and minimum free energies of GH1 mRNA-16S rRNA interactions were predicted using MFOLD (GCG-Wisconsin Package (Ver. 8.1)). Minimum free energies of predicted structures were also manually calculated, incorporating an added penalty for asymmetric internal loops (de Smit & van Duin, 1994b).

3.2.6 Isolation of bacterial ribosomes

Selected colonies were grown in 2X YT medium containing 50 μg/ml ampicillin. Overnight cultures were diluted 1:50 into fresh 2X YT medium. When the OD_{600} of these cultures reached 0.4, 0.75 mM IPTG was added to induce the expression of GH1. After a further 3 hours, cells from 3-4 ml of culture were collected by centrifugation and resuspended in lysis buffer (60 mM NH_4Cl, 60 mM KCl, 20 mM Tris-HCl pH 8.0, 6 mM $MgCl_2$, 6 mM β-mercaptoethanol and 16% sucrose). Lysis was induced by addition of 0.5 mg/ml lysozyme followed by three freeze-thaw cycles. The cell suspension was allowed to sit on ice for 30 minutes between each of the freeze-thaw cycles. Cell debris was removed by centrifugation at 13,000 rpm for 15 minutes at 4°C. The ribosome-containing lysate was transferred to a new microcentrifuge tube and stored at –70°C.

3.2.7 Sucrose gradient fraction and analysis of ribosome

Aliquots of ribosome preparation were loaded onto 10 to 30% (w/w) iso-kinetic sucrose gradients and centrifuged at 4 °C in an SW 41 rotor for $\omega^2 t = 3.7 \times 10^{11}$ $rad^2 s^{-1}$ (Liiv *et al.*, 1996). Sucrose gradients were fractionated by upward displacement and continuously analysed with a Gilson 112 UV detector. RNA from collected fractions

were purified by either extracting once with buffered phenol, once with isoamyl alcohol/chloroform, and precipitating with ethanol, or by washing through a QiaRNeasy column (Qiagen) as instructed by the manufacturer. The purified RNAs were stored at –20 °C or analysed immediately.

To analyse rRNA, an aliquot of each RNA gradient fraction was added to RNA gel loading buffer (50 mM Tris pH 8.0, 2 mM EDTA, 1% β-mercaptoethanol, 1% SDS and 10% glycerol) and were electrophoresed in a 1.4% agarose gel.

To analyse the mRNA content of ribosomes, RNA from gradient fractions were subject to Northern blotting as described above.

3.2.8 Green fluorescence protein (GFP) constructs

An expression construct containing the *Aequorea victoria* green fluorescence protein (GFP) gene was obtained from Cormack *et al.* (1996). The GFP coding sequence was removed from this construct (pKEN2) by partial restriction enzyme digestion with NdeI and HindIII. Isolated DNA was cloned into the T7 polymerase-based expression vector pT79 (Figure 3-4). The resulting construct, pGFP, coded for 238 amino acids of the green fluorescence protein. This gene had been mutated [Phe(64)-Ser(65) to Leu(64)-Thr(65)] at chromophore to increase the fluorescence (Cormack *et al.*, 1996).

Three double stranded oligonucleotides GH1-WT, GH1-3P and GH1-HP (Figure 3-5) were individually ligated into partially, NdeI-cut pGFP (Figure 3-4). Individual colonies were screened by PCR to select the desired constructs, and then after transformation into TG1 host cells, DNAs from numerous individual colonies were sequenced by the dideoxy chain-termination method in order to confirm the 3 required GFPGH1 coding sequence variants (Figure 3-6).

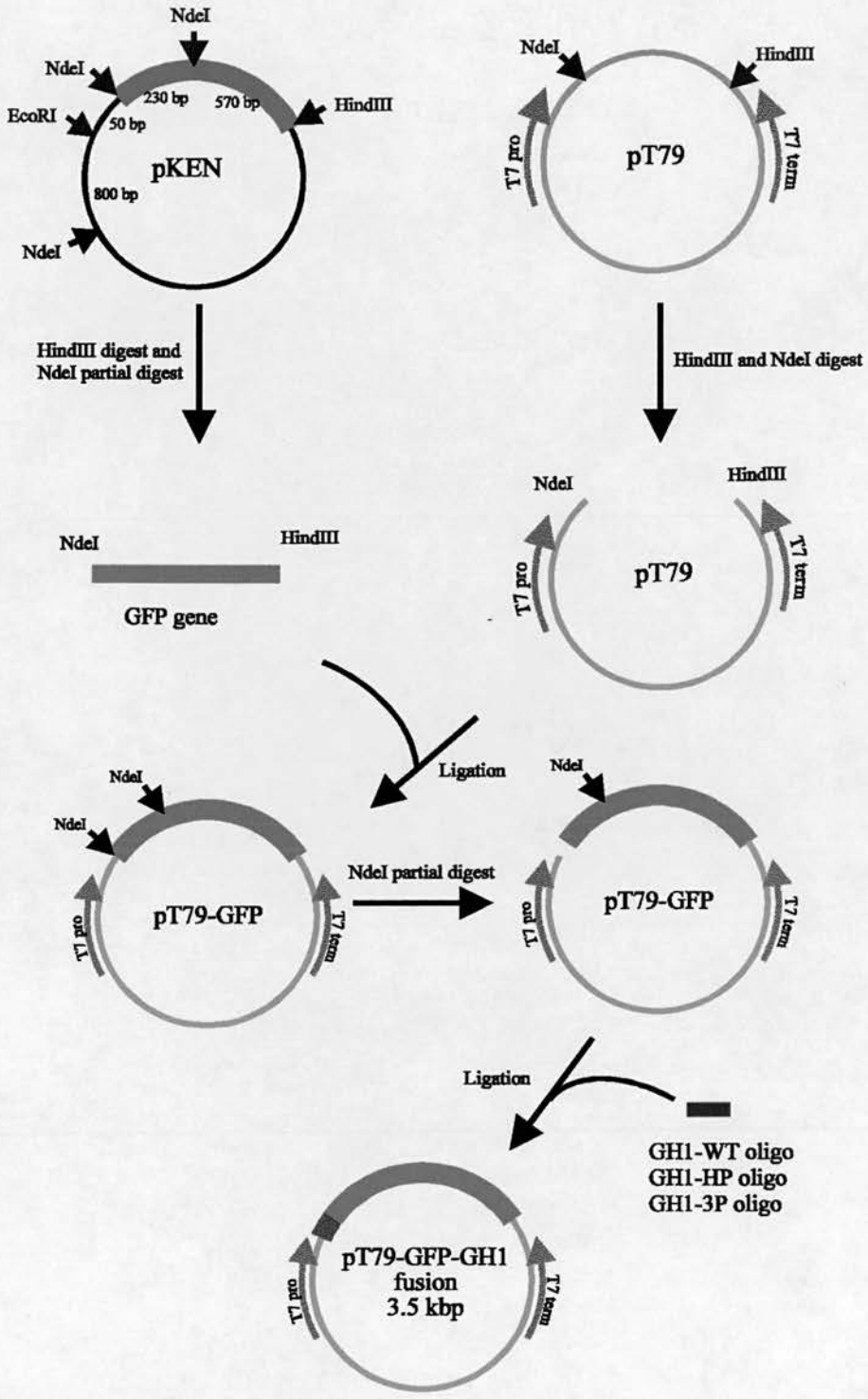


Figure 3-4 Strategy to make fusion GFPGH1 plasmids

The first 6 codons of the GFP gene in the pT79GFP

```
5'-T ATG AGT AAA GGA GAA GAA CTT
     TAC TCA TTT CCT CTT CTT GAA A-5'
```

Oligonucleotide inserts

1. GH1WT

```
5'-T ATG CCC GCG GGC CCC AGC GT
       AC GGG CGC CCG GGG TCG CAA T-5'
```

2. GH13P

```
5'-T ATG CCA GCG GGT CCA AGC GT
       AC GGT CGC CCA GGT TCG CAA T-5'
```

3. GH1HP

```
5'-T ATG CCC GCG GGG GCC AGC GT
       AC GGG CGC CCC CGG TCG CAA T-5'
```

Figure 3-5 Oligonucleotides used to insert GH1 sequence at the beginning of the GFP gene in pT79GFP.

aaat*taatac* *gactcactat* *agggag*acca caacggtttc cctctagaaa

taattttgtt taactttaag aaggagatat aca***TATGCCC*** ***GCGGGCCCCA***

GCGTT**ATG**AG TAAAGGAGAA GAACTTTTCA CTGGAGTGGT CCCAGTTCTT

GTTGAATTAG ATGGCGATGT TAATGGGCAA AAATTCTCTG TCAGTGGAGA

GGGTGAAGGT GATGCAACAT ACGGAAAACT TACCCTTAAT TTTATTTGCA

CTACTGGGAA GCTACCTGTT CCATGGCCAA CACTTGTCAC TACTCTGACT

TATGGTGTTC AATGCTTCTC AAGATACCCA GATCATATGA AACAGCATGA

CTTTTTCAAG AGTGCCATGC CCGAAGGTTA TGTACAGGAA AGAACTATAT

TTTACAAAGA TGACGGGAAC TACAAGACAC GTGCTGAAGT CAAGTTTGAA

GGTGATACCC TTGTTAATAG AATCGAGTTA AAAGGTATTG ATTTTAAAGA

AGATGGAAAC ATTCTTGGAC ACAAAATGGA ATACAACTAT AACTCACATA

ATGTATACAT CATGGGAGAC AAACCAAAGA ATGGCATCAA AGTTAACTTC

AAAATTAGAC ACAACATTAA AGATGGAAGC GTTCAATTAG CAGACCATTA

TCAACAAAAT ACTCCAATTG GCGATGGCCC TGTCCTTTTA CCAGACAACC

ATTACCTGTC CACACAATCT GCCCTTTCCA AAGATCCCAA CGAAAAGAGA

GATCACATGA TCCTTCTTGA GTTTGTAACA GCTGCTAGGA TTACACATGG

CATGGATGAA CTATACAAAT AACTGCAGGC ATGCAAGCTT aggcctctag

tctagactag aattccgatc cggctgctaa caaagcccga aaggaagctg

agttggctgc tgcca*ccgct* *gagcaataac* *tagc*

Figure 3-6 Coding sequence of the GFPGH1-WT fusion construct. (Bold italic letters indicate the GH1-WT oligo insert and italic letters indicate T7 pro or T7 term sequences.) The start codon, ATG, is underlined.

3.2.9 Expression of GFPGH1 fusion proteins

Each of the GFP fusion constructs, plus constructs pT79 and pGFP, were transformed into pLysS (BL-21) cells. Fresh single colonies were selected and then grown in L broth. Overnight cultures were diluted 1:50 into fresh LB and again grown until the OD_{600} of these cultures reached 1.3. IPTG was added to 0.75mM to induce the expression of the GFPGH1 fusion products. After a further 3 hours, the OD_{600} was again measured and then cells from 100 µl of culture were collected by centrifugation and lysed by resuspending in SDS-gel loading buffer. Total protein from equivalent amounts (OD_{600}) of culture were analysed in 15% SDS-polyacrylamide gels. Coomassie blue stained gel were scanned on a Chromoscan densitometer and the data analysed using the Bio-Rad Molecular Analyst software package. GFPGH1 expression levels were normalised by reference to a number of *E.coli* host protein bands.

3.2.10 Determination of fluorescence intensity in expression culture

For the detection of fluorescence intensity, 100 µl of a 3 hour-induced culture was collected, diluted 1:10 in phosphate-buffered saline and analysed in the Perkin Elmer 3000 Fluorescence Spectrometer. Excitation spectra was read at 488 nm and emission spectra was set at 507 nm (Cormack *et al.*, 1996).

3.2.11 *In vitro* transcription/translation

The TNT T7 Quick Coupled Transcription/Translation system (Qiagen) was used for *in vitro* transcription/translation experiments. Template plasmid DNAs were purified using the Qiaprep plasmid preparation system and further cleaned by extracting once with buffered phenol and once with iso-amylalcohol/chloroform, and precipitating with ethanol. 250 ng of plasmid DNA and 1 µl of L-[^{35}S]methionine (1000Ci/mM; Amersham) were added to 10 µl of the TNT T7 quick master mix. The reaction was performed at 30°C for 90 minutes. To analyse the results of translation, 4

µl of the reaction mix were added to SDS-gel loading buffer and loaded onto a 15% SDS-polyacrylamide gel. After electrophoresis, the radioactive products were fixed by washing the gel in 10% acetic acid and 10% methanol buffer for 10 minutes. The gel was then dried and exposed to X-ray film (Fuji).

3.3 Results

The expression of the globular domain of linker histone H1 (GH1) in *E.coli* is very poor (Gerchman *et al.*, 1994). However, it could be overcome by mutating the 3rd position of the first four codons of the GH1 coding sequence (Buckle & Allan, unpublished). With the aim of elucidating the nature of this difficulty associated with the wild-type GH1 expression construct, 16 GH1-expression constructs comprising all possible combinations of the wild-type and mutated codons were generated. The sequence of these variants are shown in Table 3-1. The plasmid vector, shared by all the constructs, was based on the tightly regulated T7 expression system (Studier, 1990) which, upon induction by IPTG, can provide abundant mRNA with optimised translation initiation regions designed for efficient, high level production of protein.

3.3.1 Expression of Recombinant GH1

3.3.1.1 Site-directed mutagenesis of GH1

Site-directed mutagenesis, using degenerate primers (Figure 3-3A), was carried out using a PCR protocol and this population of PCR products cloned into pT79 (Figure 3-3B). Plasmid DNA was isolated from 96 DE3 transformants. The GH1 insert in each of these plasmids was sequenced by the dideoxy chain termination method. An example of the sequencing results is shown in Figure 3-7. Plasmids containing the desired GH1 insert were selected and were further analysed for their expression capacity in *E.coli*.

3.3.1.2 Analysis of recombinant GH1 expression

Figure 3-8 shows an example of a 15% polyacrylamide gel analysis of total *E.coli* cell protein obtained after three hours IPTG induction of DE3 transformants harbouring each of the 16 GH1 expression plasmids. In some of lanes (notably 7, 10 and 13-17), a protein of about 8 Kd is prominent. The size of this product is consistent with the expression of GH1 (77 amino acids) and the protein is not detected in cells

Construct	DNA Sequence	Free energy (kcal/mol)	
		(a)	(b)
GH1-WT	ATGCCCGCGGGCCCC	-8.1	-9.9
GH1-11	ATGCCCGCGGGCCC**A**	-6.0	-8.2
GH1-12	ATGCCCGCGGG**T**CCC	-7.2	-10.6
GH1-13	ATGCCCGC**A**GGCCCC	-8.1	-9.1
GH1-14	ATGCC**A**GCGGGCCCC	-9.0	-10.8
GH1-21	ATGCCCGCGGG**T**CC**A**	-5.3	-8.8
GH1-22	ATGCCCGC**A**GGCCC**A**	-6.0	-8.2
GH1-23	ATGCC**A**GCGGGCCC**A**	-6.9	-8.5
GH1-24	ATGCCCGC**A**GG**T**CCC	-7.4	-10.6
GH1-25	ATGCC**A**GCGGG**T**CCC	-7.8	-11.5
GH1-26	ATGCC**A**GC**A**GGCCCC	-9.0	-9.5
GH1-31	ATGCCCGC**A**GG**T**CC**A**	-5.5	-8.8
GH1-32	ATGCC**A**GCGGG**T**CC**A**	-7.8	-9.7
GH1-33	ATGCC**A**GC**A**GGCCC**A**	-6.9	-8.1
GH1-34	ATGCC**A**GC**A**GG**T**CCC	-7.8	-11.5
GH1-41	ATGCC**A**GC**A**GG**T**CC**A**	-7.8	-9.7

Table 3-1 The name of 16 constructs which were made by site-directed mutagenesis were shown and followed by its first 4 codons. Free energy was calculated by MFOLD in (a) 80 nucleotides (-43 to +37) and (b) 90 nucleotides (-43 to +47).

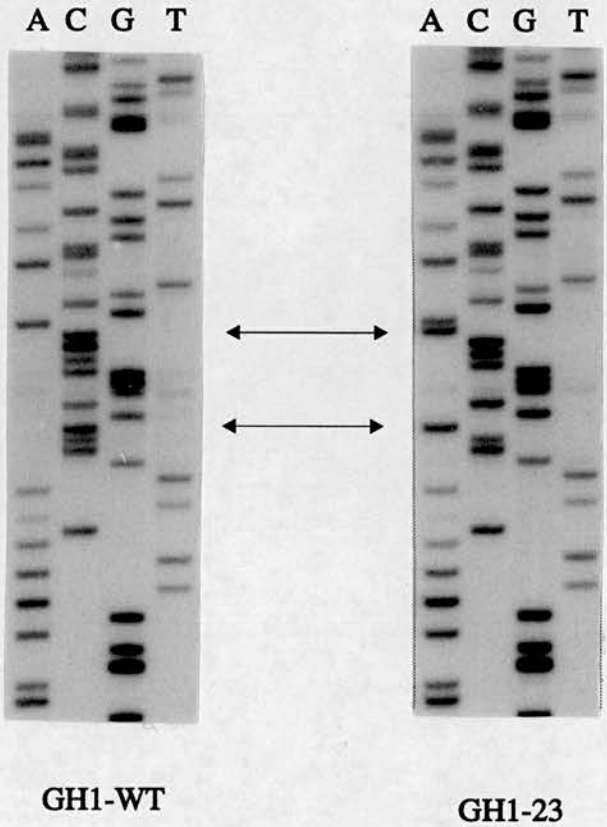


Figure 3-7 DNA sequence of GH1-WT and one of the GH1 mutants (GH1-23). In the mutant, both the first and fourth codons have been changed from CCC to CCA (arrowed).

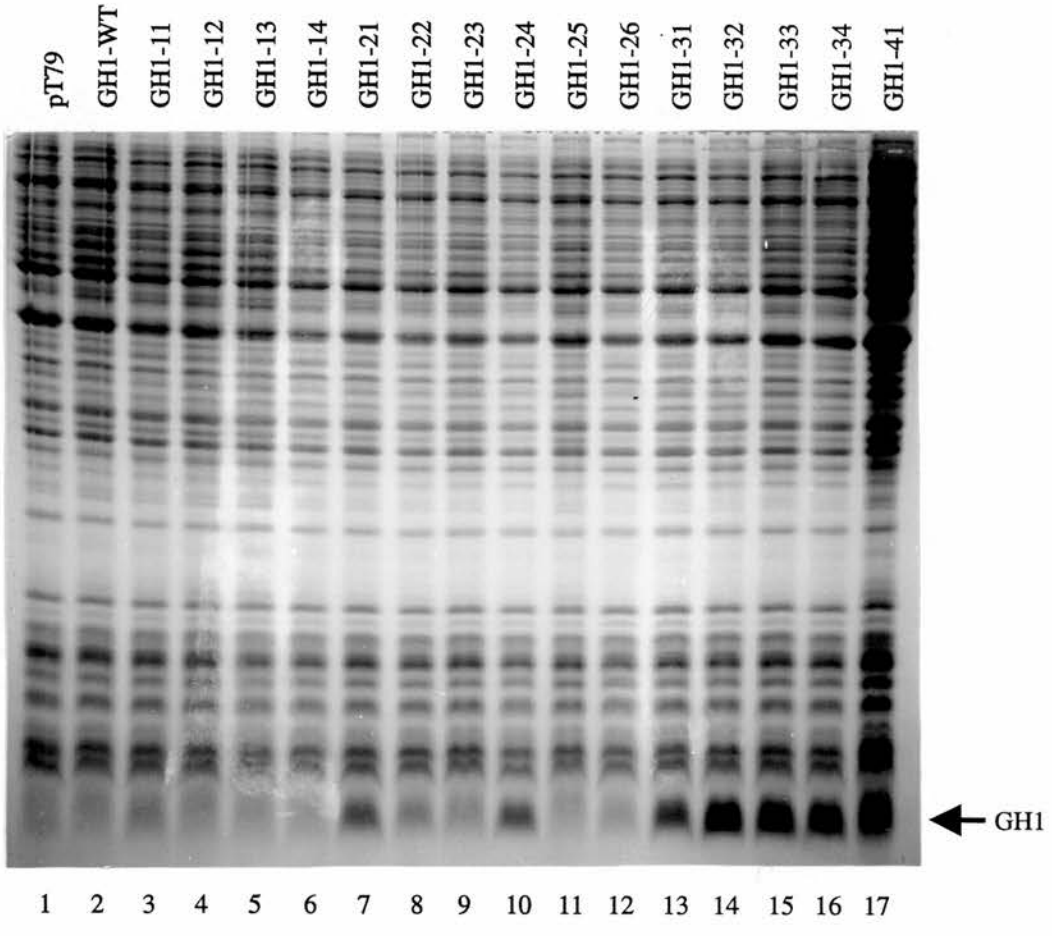


Figure 3-8 15% SDS-polyacrylamide gel analysis of total proteins isolated from IPTG-induced DE3 cells harbouring each of the GH1 expression constructs. Proteins from a vector-only transformant (pT79) are shown in Lane 1. The position of GH1 is indicated.

transformed with an expression vector lacking a GH1 insert (lane 1). This 8 Kd protein is uniquely soluble in perchloric acid (see Figure 4-7), a property consistent with its identification as GH1 (Allan *et al.*, 1980). Furthermore, the purified peptide displays chromatosome protection in a reconstitution assay (see Figure 4-8), a property which is diagnostic for a functional linker histone globular domain (Allan *et al.*, 1980). [Chromatosome protection involves reconstituting linker histone-depleted polynucleosomes with H1 or GH1 (recombinant GH1), and digesting the resulting chromatin with micrococcal nuclease. In the absence of linker histone, only 146 bp of DNA is protected from nuclease digestion, whereas the presence of correctly bound H1/GH1 a further 20 bp of DNA is protected (168 bp).]

In the context of acid solubility and chromatosome protection, the 8 Kd product was taken to be recombinant GH1 and to be functionally equivalent to its native counterpart.

A comparison of lanes 1 and 2 reveals that the expression of GH1 from the wild-type sequence is barely detectable. Similarly, all of the one-point mutants (lanes 3-6) are poorly expressed. Two of the two-point mutants, GH1-21 (lane 7) and GH1-24 (lane 10) show significant expression while all of the three-point and four-point mutants display high levels of GH1 expression. Coomassie stained polyacrylamide gels from three independent experiments were quantified by densitometry and a summary of the results is shown in Figure 3-9. The maximum enhancement of GH1 expression brought about by mutation, relative to the amount obtained with the wild-type GH1 plasmid, is about 300-fold.

These results indicate a correlation between the number of codons mutated and the level of GH1 expressed and suggest that the expression of GH1 in *E.coli* is influenced by the overall AT content of the initial codons. Many factors could contribute to the diversity of expression, such as the levels of GH1 mRNA, codon usage or mRNA secondary structure. Further analyses were undertaken to clarify this issue.

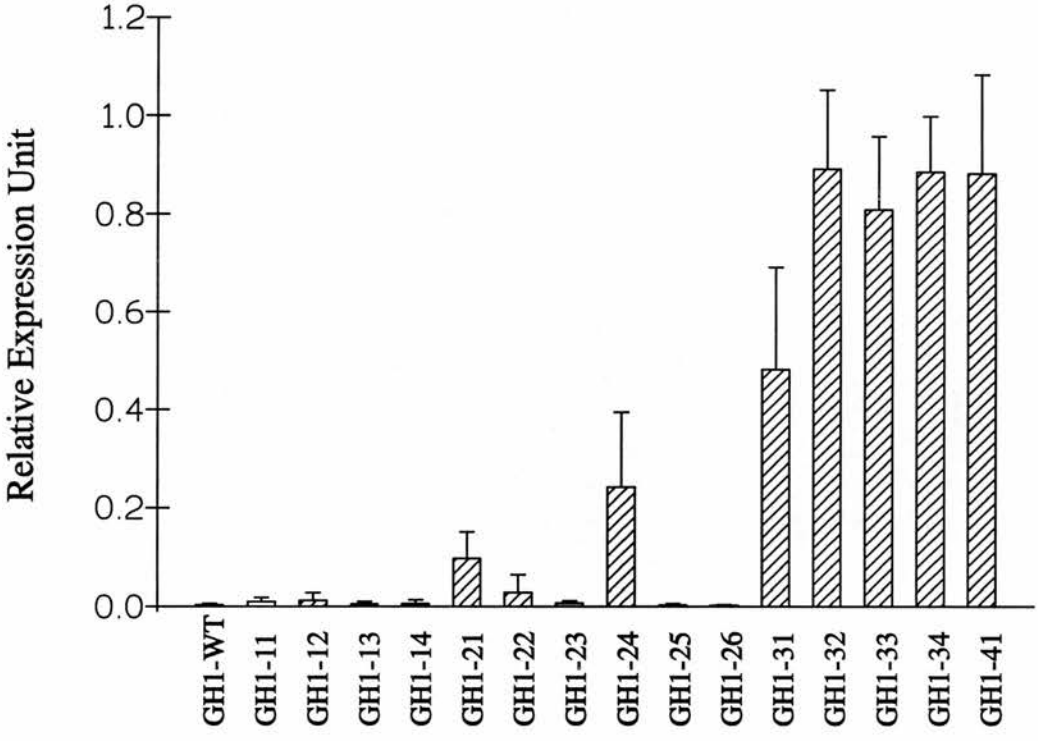


Figure 3-9 Quantitative evaluation of GH1 synthesis for each expression construct. The data represents an average of three independent experiments.

3.3.2 Amount of GH1 mRNA transcript

From the expression studies, it was shown that there is a diversity of expression levels between the 16 GH1 expression constructs. This result may reflect a differing ability to produce mRNA or from differences in mRNA stability. Therefore, Northern-blot hybridisation was employed to assess whether variation in the amount of GH1 protein produced by the different expression constructs (Figure 3-9) reflected the amount of available mRNA transcript.

Fresh cell pellets, collected after three hours IPTG induction of DE3 transformants harbouring each of the GH1 expression plasmids, were resuspended in RNA sample buffer, and fractionated on a formaldehyde-containing agarose gel. After transfer to a nitrocellulose filter, α-dCTP^{32} labelled, wild-type GH1 DNA was employed to probe the filter for GH1 mRNA. As shown in Figure 3-10, two major bands were detected. Identification of the correct GH1 mRNA transcript was determined by using an *in vitro* transcription assay. The mCAP mRNA capping kit (Stratagene) was employed to synthesise wild-type GH1 mRNA using the T7 RNA polymerase promoter. This approach identified the upper band (Figure 3-10) as the correct GH1 mRNA transcript (~ 246 nucleotides) (data not shown). The lower band may represent degradation products of the full length GH1 mRNA.

A comparison between the levels of GH1 mRNA present in extracts shows that they are similar for all constructs (Figure 3-10). Minor differences in mRNA levels were not consistent between independent analyses and did not correlate with GH1 protein expression. From Figure 3-10, for example, mRNA levels of GH1-WT are more abundance than mRNA levels of GH1-41. The results of this analysis, therefore, exclude the possibility that expression diversity reflects the level of GH1 mRNA derived from the different constructs.

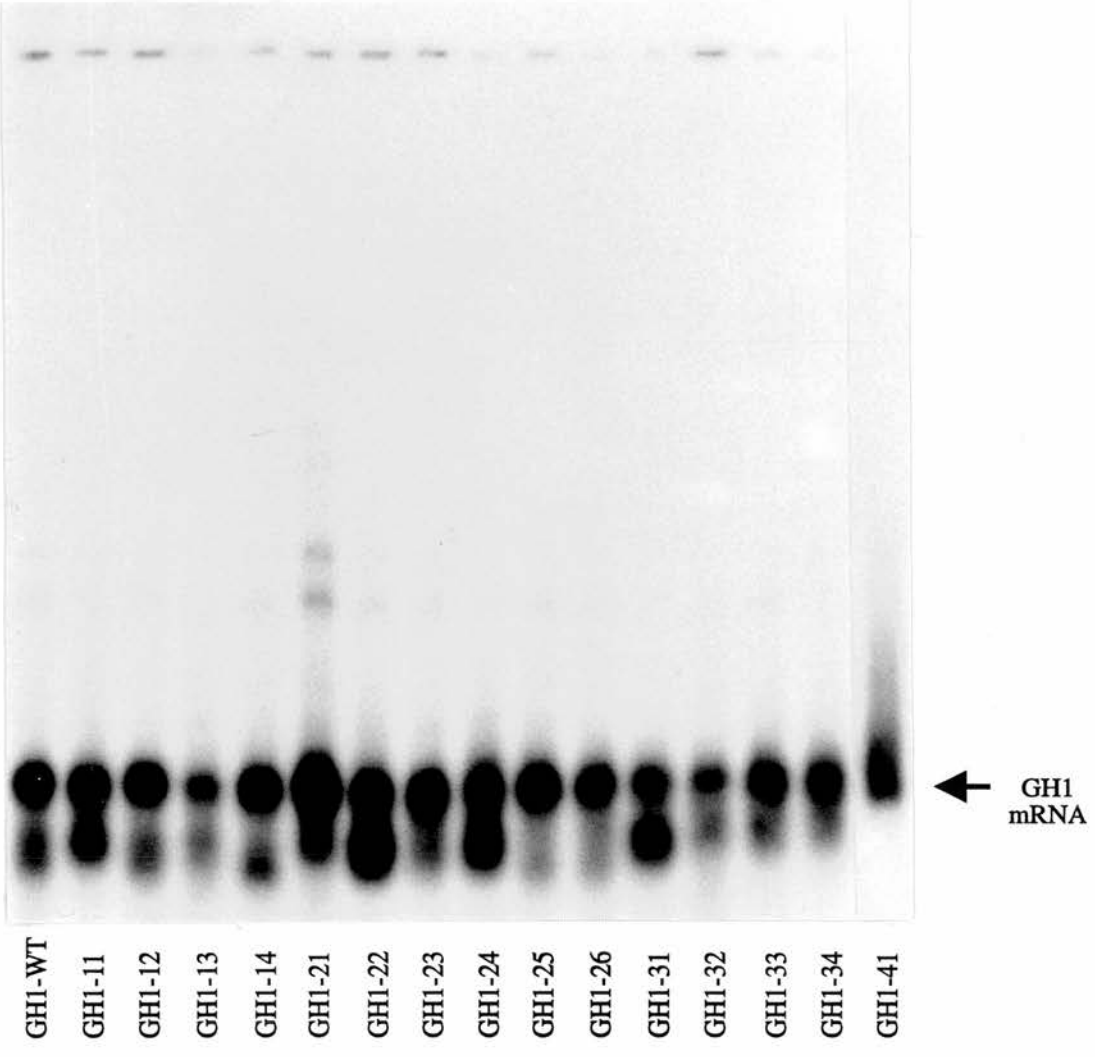
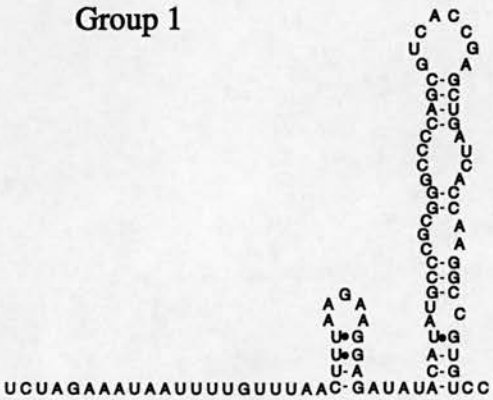


Figure 3-10 Northern blot analysis of GH1 mRNA levels in extracts from IPTG-induced DE3 cells harbouring each of the GH1 expression constructs.

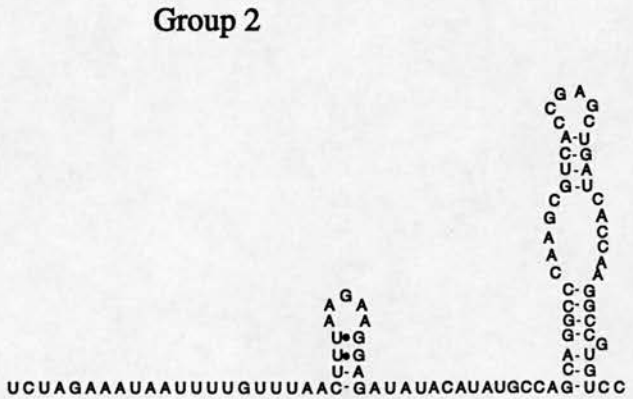
3.3.3 Prediction of GH1 mRNA secondary structure

Efficient initiation of translation requires that the Shine-Dalgarno sequence at the 5′ end of the mRNA is accessible to base-pair with its complementary sequence at the 3′ end of the 16S rRNA. This accessibility is often an inverse function of the capacity of the mRNA to form secondary structure in the region of the Shine-Dalgarno sequence (de Smit & van Duin, 1990; de Smit & van Duin, 1994a; de Smit & van Duin, 1994b). MFOLD (GCG-Wisconsin package (Ver. 8.1)) was employed to predict mRNA secondary structure and to calculate minimum free energies for each of the 16 GH1 transcripts. This analysis was applied to a section of the mRNA comprising 43 nucleotides of leader sequence 5′ of the AUG and either 37 or 47 nucleotides of the coding sequence. The leading 43 nucleotides contains the ribosome-binding site and includes the Shine-Dalgarno sequence. The predicted secondary structure for this section of the sequence was identical for each of the GH1 mRNAs. The predicted secondary structure of the 16 mRNAs did however differ with sequence within the GH1 coding sequence regions (Figure 3-11). For the 80 and 90 nucleotide analyses, structures were distributed between 6 or 4 variants respectively, and displayed a limited spread in minimum free energy. Clearly, changes of only a few nucleotides does not affect the minimum free energy of the mRNA structure very much (Table 3-1). Furthermore, there was no correlation between translational efficiency and the stability of predicted mRNA structure (Figure 3-12). For example, in the 80 nucleotide analysis (Table 3-1), structures predicted for the single-point mutants, which as a group expressed very poorly (Figure 3-9), had an average minimum free energy of -7.6 kcal/mol whereas for the 3 and 4 point mutants, which expressed well, the corresponding value was -7.2 kcal/mol.

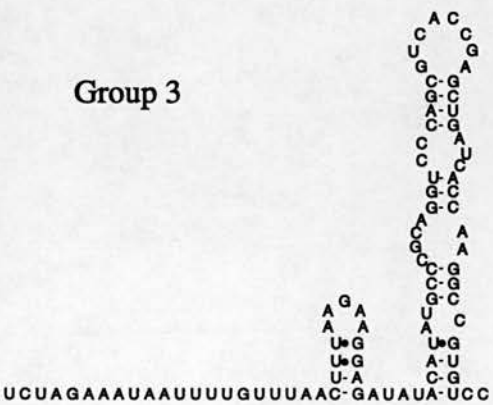
The minimum free energies for predicted secondary structures were also calculated according to the method of de Smit & van Duin (1994b), incorporating an added penalty for asymmetric internal loops, but again no correlation between mRNA structure stability and GH1 expression occurred (Figure 3-12). These observations may indicate that the control of GH1 gene expression is not simply a function of mRNA secondary structure stability.



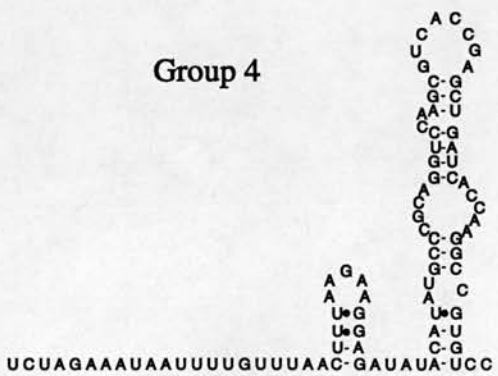
GH1-WT
GH1-11
GH1-13
GH1-14
GH1-22
GH1-23
GH1-26



GH1-33



GH1-12
GH1-24
GH1-25
GH1-34



GH1-21
GH1-31
GH1-32
GH1-41

Figure 3-11 Prediction of GH1 mRNA secondary structure for 90 nucleotides analysis. In group 1, the structure of GH1-WT was shown. In group 2, the structure of GH1-33 was shown. In group 3, the structure of GH1-24 was shown. In group 4, the structure of GH1-31 was shown.

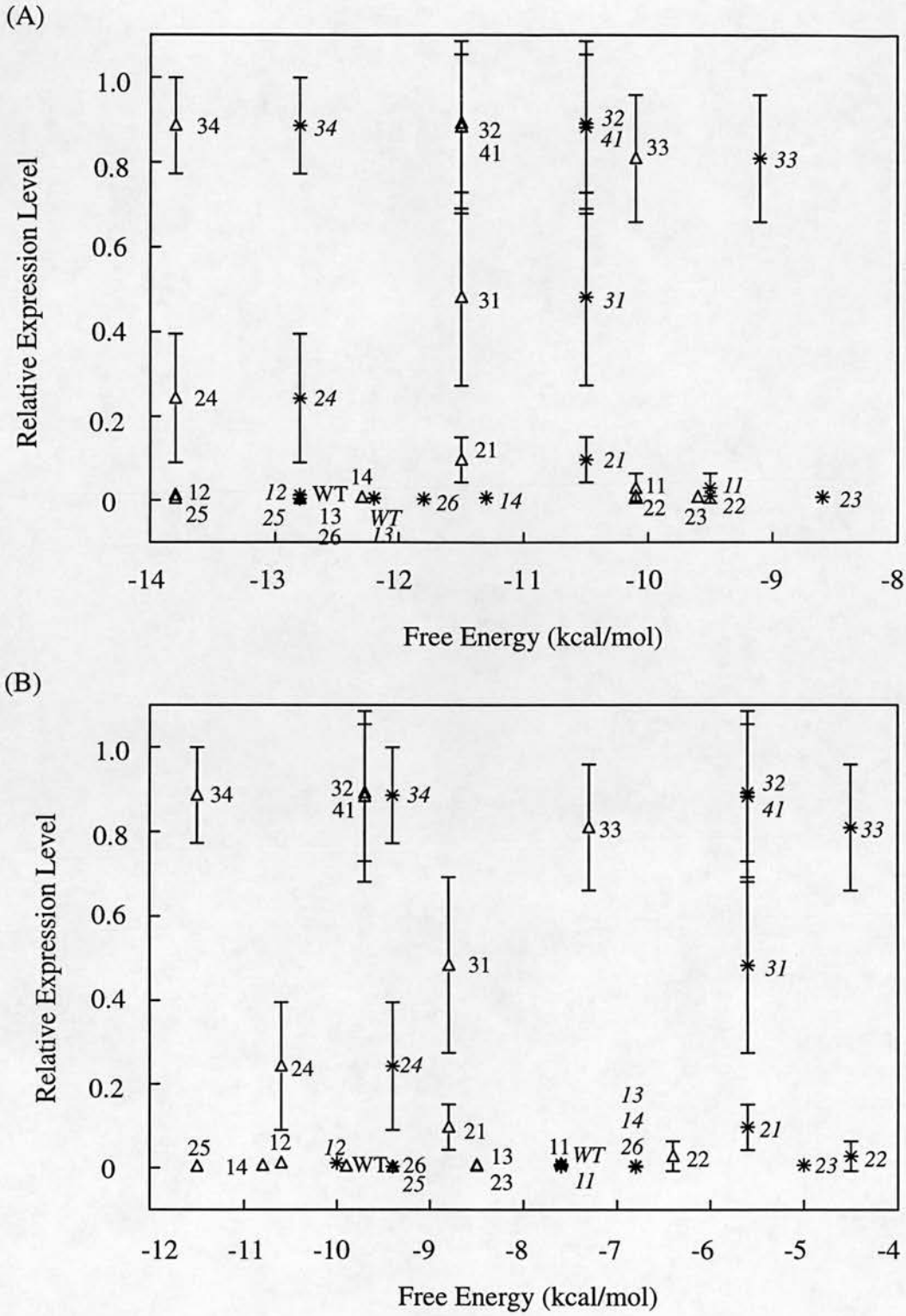


Figure 3-12 Correlation between the expression of GH1 (from Figure 3-9) and the stability of mRNA secondary structure calculated by either (A) FOLDRNA or (B) MFOLD (GCG-Wisconsin Package (Ver. 8.1)) for 90 nucleotides (-43 to +47). (△: the free energy value predicted from GCG-Wisconsin package). (✱: the free energy value for predicted secondary structures according to the method of de Smit & van Duin (1994a)). Each sample is marked either in normal (△) or in Italic form (✱) number. The MFOLD values predicted from GCG-Wisconsin package are listed in Table 3-1.

3.3.4 Association of GH1 mRNA with ribosome subunits

The translation initiation region (TIR) in *E.coli* mRNA comprises the start codon and the Shine-Dalgarno region (Shine & Dalgarno, 1975). However, these domains are not always sufficient to define an efficient TIR (Stormo *et al.*, 1982). It has been suggested that additional sequences (or structures) may interact with the 16S rRNA to modulate expression (see Chapter 3.1.3.3; Sprengart *et al.*, 1990; Faxen *et al.*, 1991; Nagai *et al.*, 1991; Ito *et al.*, 1993; Sprengart *et al.*, 1996). How translation is affected by mRNA sequences upstream of the Shine-Dalgarno region or downstream of the start codon is not clear. GH1 expression has shown that minor changes in the first 4 codons can have profound effects on GH1 expression (for example, the GH1-23 and GH1-32 only differ in single nucleotide, but they have ~300-fold differences in expression (Table 3-1, Figure 3-9), and yet there is no clear correlation between mRNA structure stability and GH1 expression. It may be possible that the interaction of GH1 mRNA with the 30S ribosomal subunit has been stalled at some point during the translation initiation process such that the complex cannot join efficiently with the 50S component to form a productive 70S initiation complex and activate translation elongation (Figure 3-1). To investigate this possibility, an analysis of the distribution of GH1 mRNA between ribosomal components was undertaken.

Two plasmids, GH1-WT and GH1-32, were chosen for this analysis as being representative of low and high expression levels, respectively. The plasmids were grown in 2X YT medium and after 3 hours induction, cultures were harvested and ribosomes were prepared as described by Liiv *et al.* (1996). The ribosome preparations were then fractionated on 10-30% sucrose gradients to separate the different components, namely the 30S, 50S and 70S subunits. Sucrose gradient profiles from each sample are shown in Figure 3-13. Peaks corresponding to the 30S subunit, the 50S subunit and the 70S subunit are clearly separated. The identity of each component peak was confirmed by agarose gel analysis of rRNA from gradient fractions (Figure 3-14).

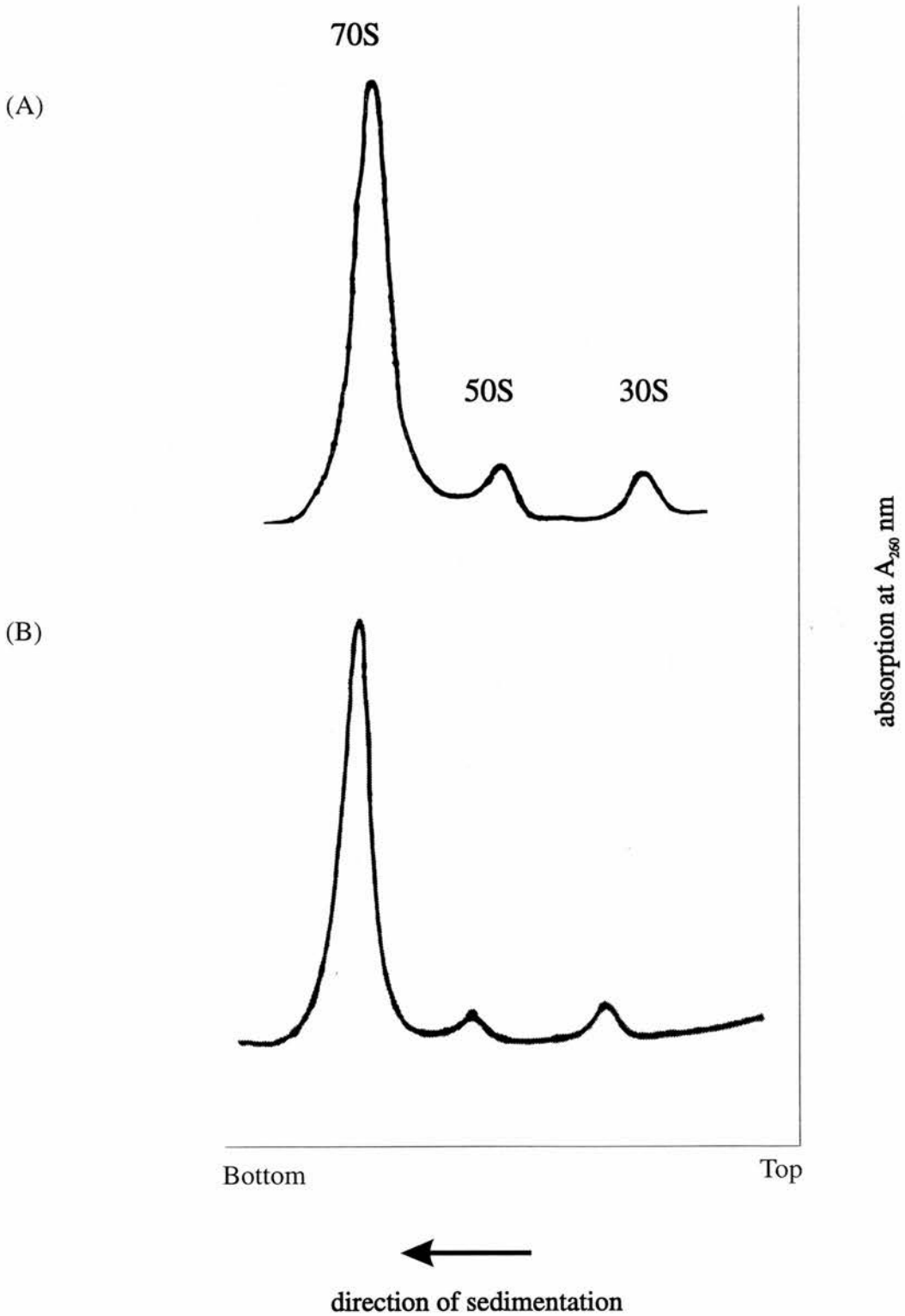


Figure 3-13 Sucrose gradient profiles of ribosomes from cells harbouring 2 different GH1 plasmids : (A) GH1-WT; (B) GH1-32.

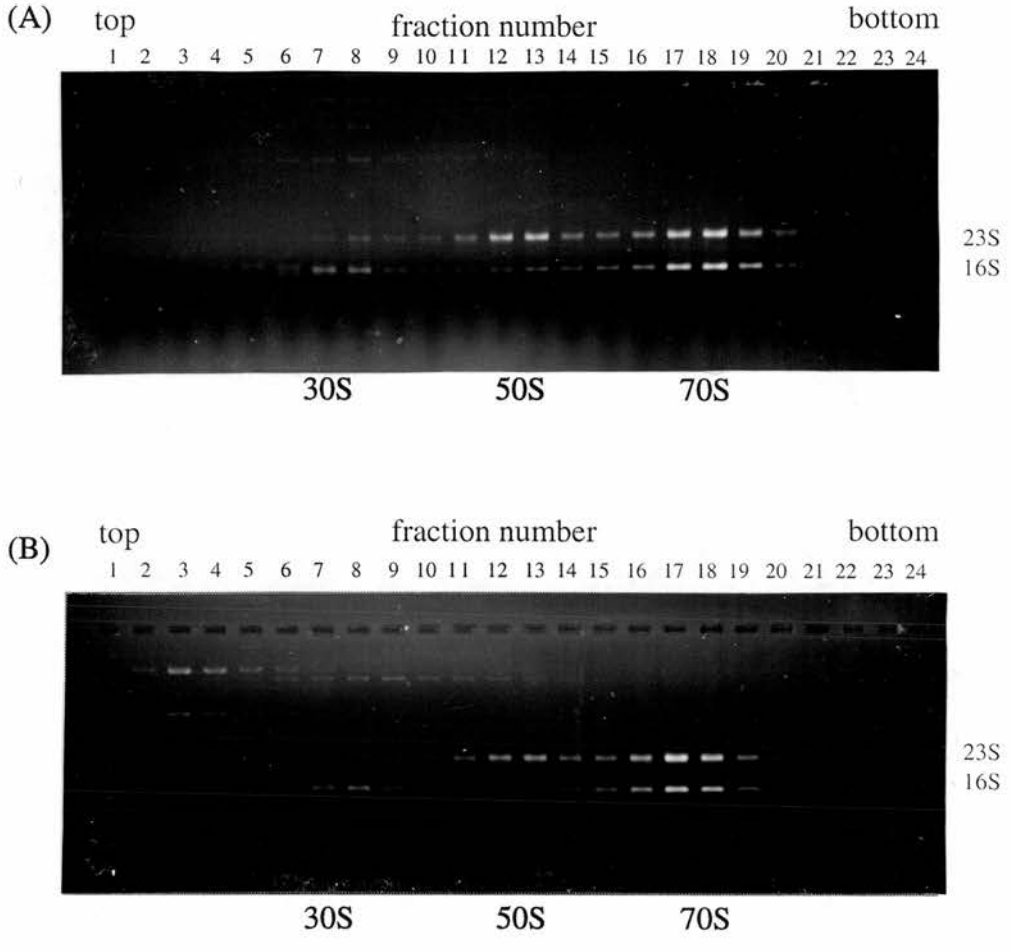


Figure 3-14 1.4% agarose gel analysis of ribosome subunits in sucrose gradient fractions from IPTG-induced DE3 cells harbouring GH1-WT (A) or GH1-32 (B) expression constructs.

To identify the GH1 mRNA contained in each subunit, Northern hybridisation was employed. Aliquots of RNA isolated from gradient fractions for both GH1-WT and GH1-32 were loaded onto a 1% formaldehyde agarose gel. After electrophoresis and transfer to nitrocellulose, the filter was hybridised with an α-dCTP^{32} labelled GH1-WT probe. The results of this analysis are shown in Figure 3-15. For the low-expressing plasmid, GH1-WT, the 30S ribosomal component is associated with substantial GH1 mRNA; particularly no mRNA was detected in the 50S ribosomal component and only a minor fraction of GH1-WT mRNA was seen in the 70S ribosomal component. In contrast, for the high expressing plasmid, GH1-32, the results were quite different. Here, substantial amounts of GH1 mRNA were found associated with the 70S ribosomal component, none detected in the 50S ribosomal component, and only a minor fraction of GH1 mRNA was in the 30S ribosomal component. For ease of comparison, the Northern blots were quantitated by densitometry and the results are shown in Figure3-16.

These results indicate that the transfer of GH1-WT mRNAs into a productive 70S subunit is less effective than that achieved with GH1-32 mRNA. The GH1-WT mRNA tends to accumulate in the 30S subunit, whereas GH1-32 mRNA is efficiently transferred into the 70S subunit. These experiments suggest that the two types of mRNA differ in their ability to be incorporated into a productive 70S subunit which is likely to be an important factor in determining expression levels.

3.3.5 Expression of GFP-GH1 fusion proteins

Results already obtained have shown that the first four codons of GH1 have a substantial effect on the expression level of GH1 protein. Furthermore, this may correlate to the efficiency with which the mRNA is transferred into a productive 70S complex. To test whether this is a property inherent to the sequence of the four codons alone, these short sequences were introduced in front of another gene, which is highly-expressed from the T7 promoter in *E.coli.*

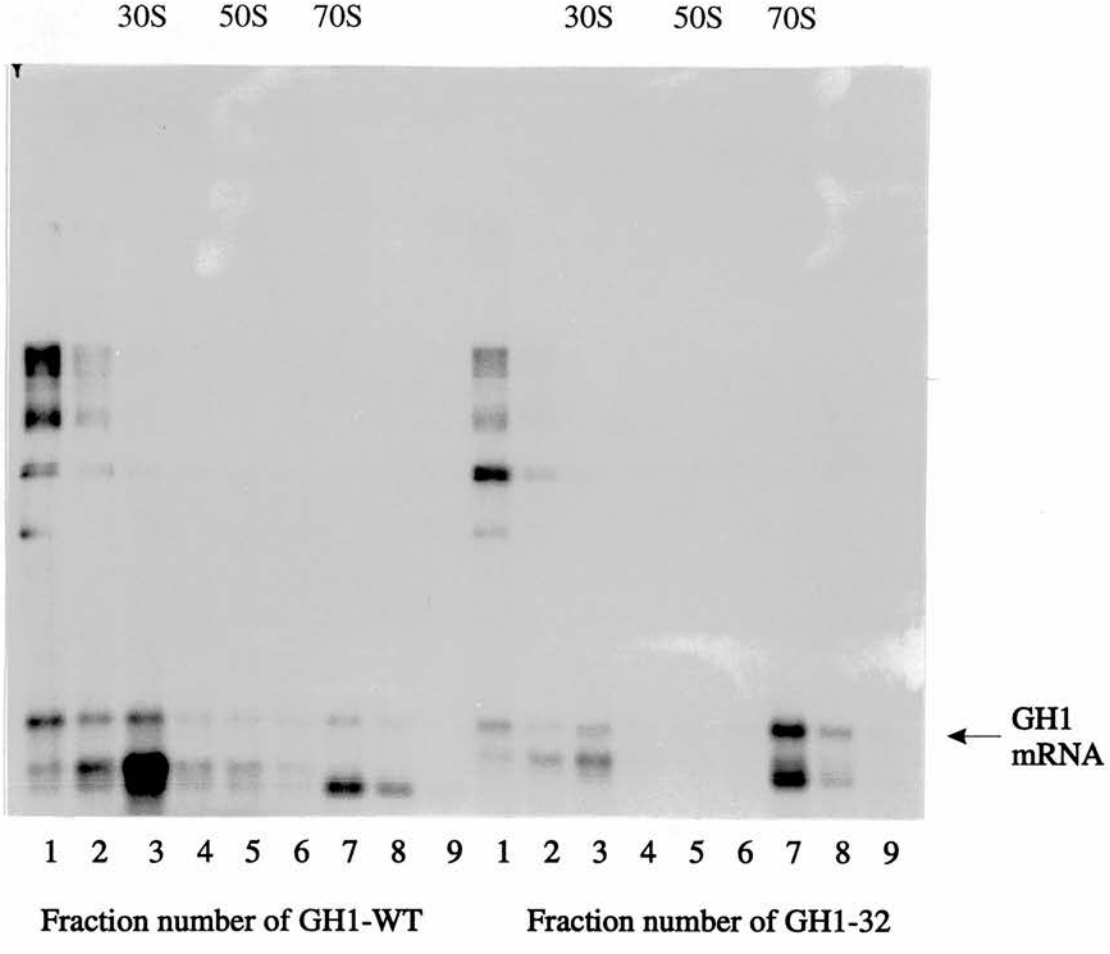


Figure 3-15 Northern blot analysis of GH1 mRNA levels in sucrose gradient fractions from IPTG-induced DE3 cells harbouring GH1-WT or GH1-32 expression constructs.

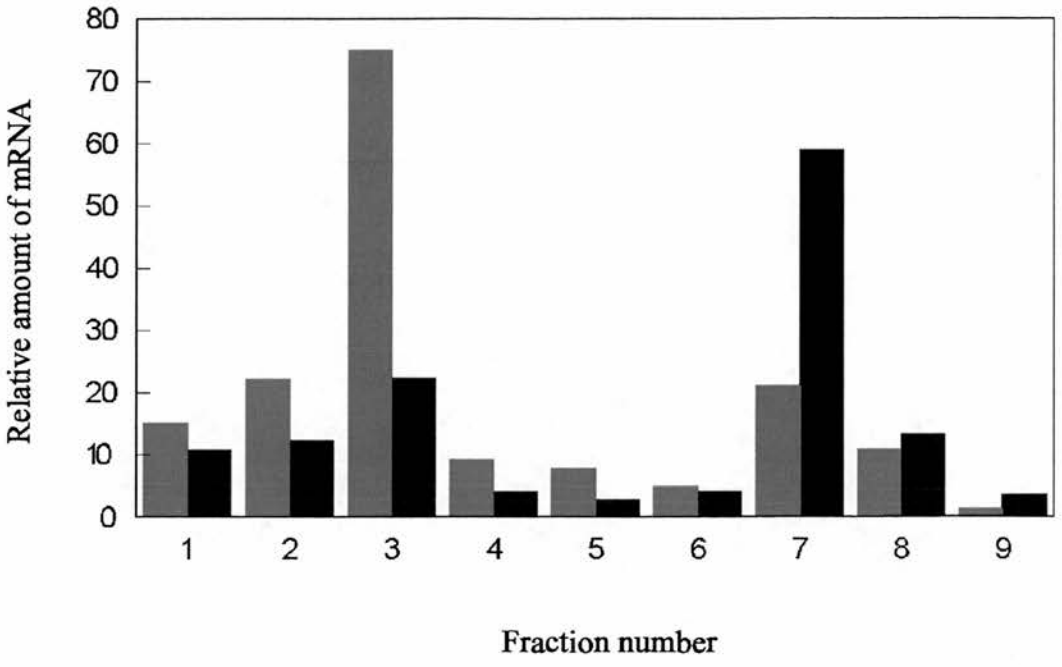


Figure 3-16 Quantitation of mRNA amount in sucrose gradient fraction for GH1-WT (red column) and GH1-32 (blue column). Fraction 3 contains 30S subunit, fraction 5 contains 50S subunit, and fraction 7 contains 70S.

The green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* (Prasher *et al.*, 1992) absorbs blue light with an excitation maximum of 395 nm, and emits green light with a maxima of 510 nm (Morise *et al.*, 1974; Ward *et al.*, 1980). This fluorescence is very stable, and virtually no photobleaching is observed. The GFP chromophore is derived from the primary amino acid sequence through the cyclization of serine-dehydrotyrosine-glycine within a hexapeptide that starts at amino acid 64 (Cody *et al.*, 1993). The mechanism that produces the dehydrotyrosine and cyclization of the polypeptide to form the chromophore are unknown. The ability to generate fluorescence *in situ* by expressing the gene for GFP has opened up tremendous possibilities for continuously monitoring gene expression in both *E.coli* (Chalfie *et al.*, 1994; Inouye & Tsuji, 1994) and in higher eukaryotes (Chalfie *et al.*, 1994; Wang & Hazelrigg, 1994). Therefore, the GFP plasmid was selected as a host expression construct to assess the assertion that the first four codons of GH1 specifically effect expression levels.

The GFP plasmid used in these studies was obtained from Cormack *et al.* (1996). In the original plasmid, pKEN, site-directed mutation had been employed to substitute Phe (64)-Ser (65) with Leu (64)-Thr (65), to enhance fluorescence 35-fold over the wild type. To make this study comparable to the GH1 work, the GFP gene was cloned into the pT79 vector to produce pT79GFP. The method used to clone the GFP gene into the pT79 vector is shown in Figure 3-4.

3.3.5.1 Strategy to express GFP-GH1 fusion

To test the hypothesis that the first few GH1 codons have great influence on gene expression, three double-stranded oligonucleotides, GH1-WT, GH1-3P and GH1-HP, were introduced into the first NdeI site in the pT79GFP plasmid. The GH1-WT oligonucleotide is the same as the first 5 codons of the plasmid GH1-WT, and the GH1-3P oligonucleotide is the same as the first 5 codons of the plasmid GH1-32. The GH1-HP oligonucleotide is described later for the purpose of discussion. Ligated plasmids were transformed into *E.coli* TG1 cells. Successful transformants were selected by plating on LB-agar medium containing ampicillin (50μg/ml). Clones that

contained the oligonucleotide insert DNA at the correct site and in the correct orientation were detected by colony PCR.

Plasmid DNA of selected correct clones were transformed into a second bacterial host, *E.coli* strain BL21 (DE3) *Lys* S. This host contains a T7 RNA polymerase gene downstream of the *lac* promoter which can be induced by IPTG. Plasmid DNAs were isolated from numerous BL21 (DE3) *Lys* Stransformants. The GFP-GH1 insert in each of these plasmids was sequenced by the dideoxy chain termination method to assure cloned sequence in final expression host. An example of the sequencing results is shown in Figure 3-17. Plasmids containing the correct GFP-GH1 insert were analysed for their expression capacity in *E.coli*.

3.3.5.2 Expression of GFP-GH1 fusion proteins

The ability of the 3 GFPGH1 plasmids and 2 control plasmids (pT79 and pT79GFP) to express GFP in *E.coli* was investigated by polyacrylamide gel analysis. Figure 3-18 shows an example of a 15% polyacrylamide gel analysis of total *E.coli* cell protein obtained after three hours IPTG induction of DE3 transformants harbouring each of the GFP-GH1 expression plasmids. In lane 4, induced pT79GFP, a protein of about 24 Kd is prominent. The size of this product is consistent with the expression of GFP and the product is not detected in cells transformed with expression vector lacking a GFP insert - pT79 (lane 2). Finally, this 24 Kd protein is uniquely visible under UV- illumination, a property consistent with its identification as GFP (Chalfie *et al.*, 1994).

A comparison of GFPGH1-WT, positive control (pT79GFP) and negative control (pT79) reveals that the expression of GFP from the GFPGH1-WT sequence is barely detectable between pre-induction and post-induction (Figure 3-18, lanes 1-6). Similarly, the GFPGH1-HP fusion (lanes 7 & 8) is poorly expressed. In contrast, GFPGH1-3P expresses substantial amounts of GFP post-induction (Figure 3-18, lanes 9 & 10). Coomassie stained polyacrylamide gels from three independent experiments were quantified by densitometry and a summary of the results is shown in Figure 3-19. The maximum enhancement of GFPGH1-3P expression brought about by the GH1-3P

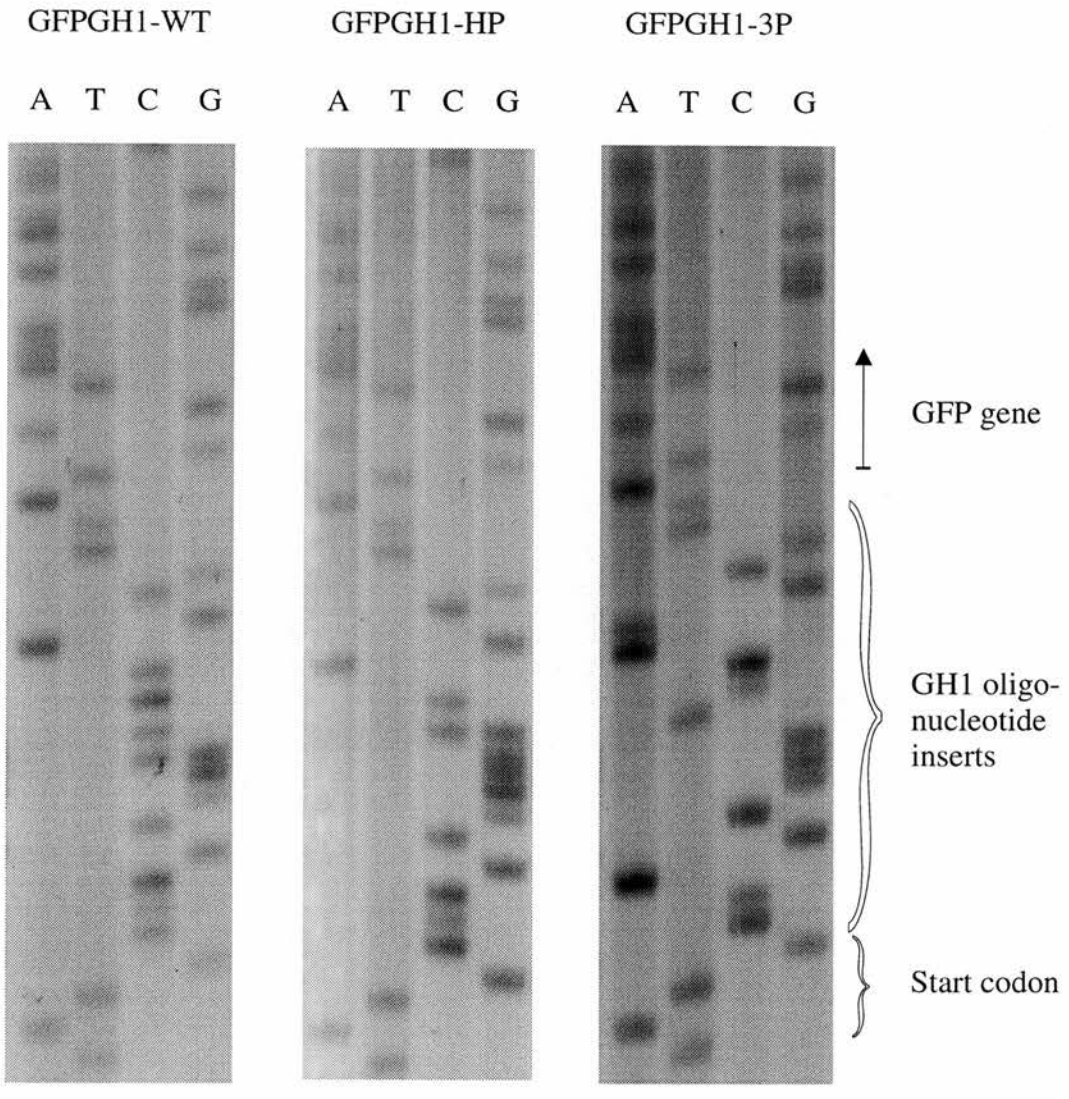


Figure 3-17 DNA sequence of GFPGH1 plasmids. The sequence of GH1WT insert is 5'-T ATG CCC GCG GGC CCC AGC GT-3', GH13P insert is 5'-T ATG CCA GCG GGT CCA AGC GT-3', and GH1HP insert is 5'-T ATG CCC GCG GGG GCC AGC GT-3'.

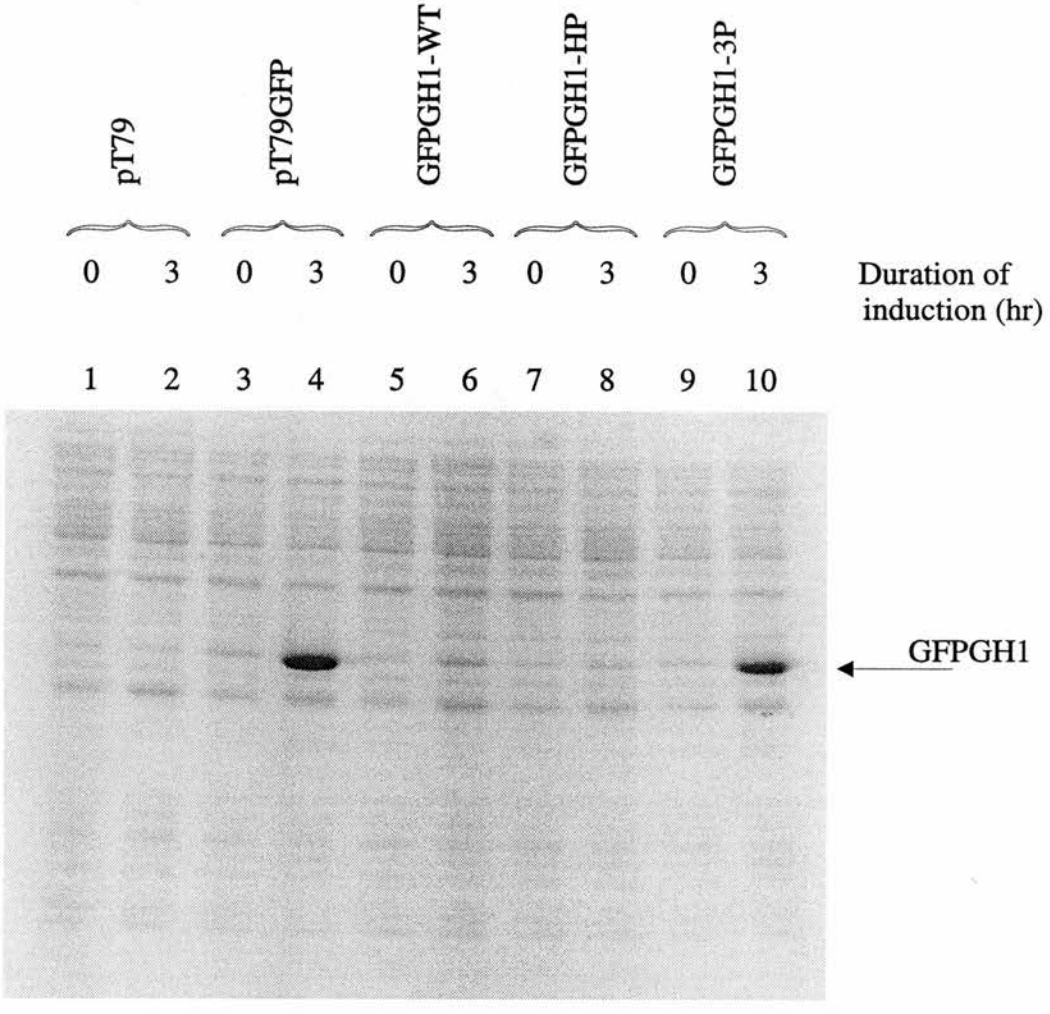


Figure 3-18 15% SDS-polyacrylamide gel analysis of total proteins isolated from IPTG-induced pLysS cells harbouring GFPGH1 expression constructs. Proteins from a vector-only transformant (pT79) are shown in Lanes 1&2. Proteins from a GFP gene-only transformant (pT79-GFP) are shown in Lanes 3& 4. The position of GFPGH1 is indicated.

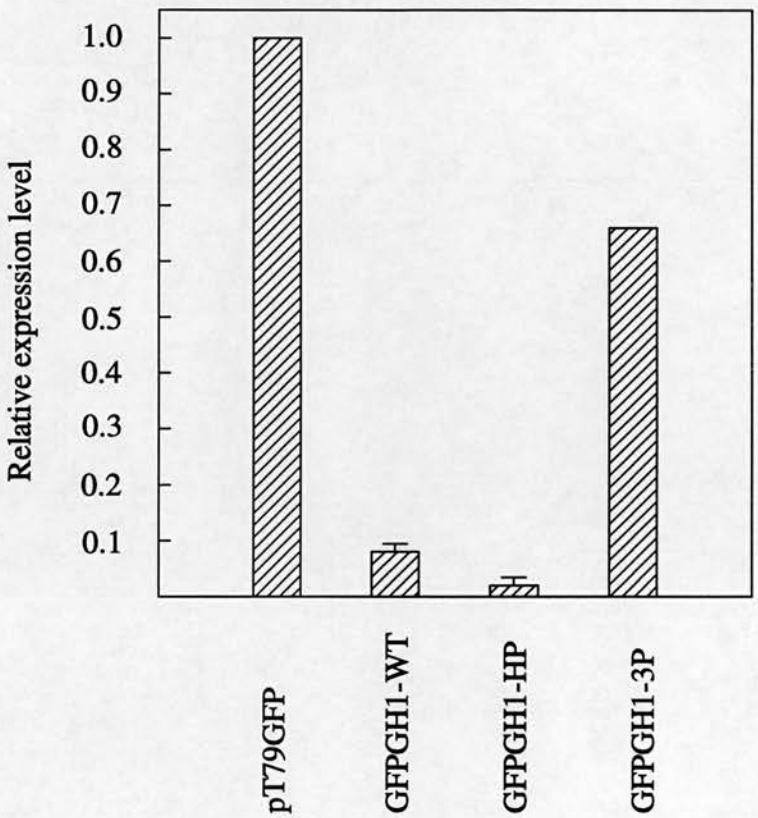


Figure 3-19 Quantitative evaluation of GFPGH1 fusion protein synthesis for each expression construct. The data represent an average of three independent experiments.

insert, relative to the amount obtained with the GFPGH1-WT is about 8.25. The results clearly demonstrate a correlation between the nature of the first few codons of GH1 and the level of GFP expression.

3.3.5.3 Fluorescence intensity of expressed GFP-GH1 fusion protein

The GFP protein can produce a fluorescent product when expressed in prokaryotes (Chalfie *et al.*, 1994; Inouye & Tsuji, 1994) or eukaryotes (Chalfie *et al.*, 1994; Wang & Hazelrigg, 1994). Because exogenous substrates and cofactors are not required for this fluorescence, fluorescence can be used to monitor gene expression. This property was also used to measure expression from the GFPGH1 plasmids.

Overnight cultures were inoculated into fresh LB. When the OD_{600} of these cultures reached 1.0, 0.75 mM IPTG was added to induce the expression of GFP. After induction with IPTG, 100 µl aliquots of expression cultures were taken at various time points to measure the fluorescence intensity when excited at 488 nm, and therefore to determine levels of GFP expression as a function of time after induction.

As shown in Figure 3-20, the positive control (pT79GFP) displays a substantial increase in fluorescence intensity within the first 30 minutes post-induction, and continuously increases in intensity to the end of time course, whereas the negative control (pT79) shows no detectable fluorescence. This result confirms that the fluorescence intensity can be used to monitor the levels of GFP expression in this system. The high-expressing plasmid, GFPGH1-3P, shows a similar fluorescence-increasing pattern as pT79GFP. In contrast, both GFPGH1-WT and GFPGH1-HP show slow increases in their fluorescence intensity post-induction and after 120 minutes induction, both plasmids have reached their fluorescence intensity plateau. The ratio of GFPGH1-3P to GFPGH1-WT fluorescence is 3.4:1 after 180 minutes induction (Figure 3-20).

This result confirms that the GFPGH1-3P construct expresses GFP much more efficiently than the other plasmids (GFPGH1-WT & GFPGH1-HP), and that this effect is modulated by the first 4 codons of the coding sequence.

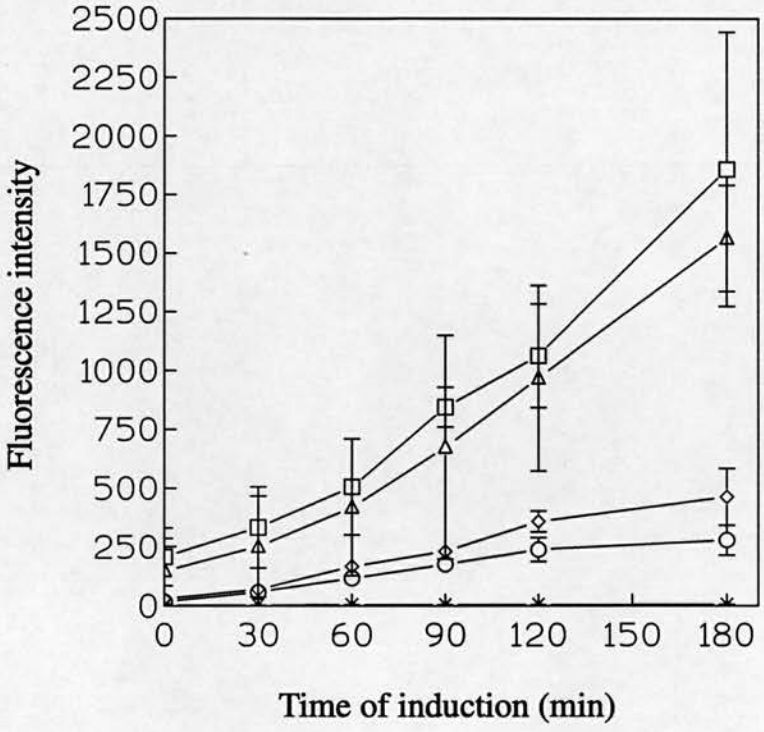


Figure 3-20 Time course of fluorenscence intensity of different constructs after IPTG induction. (∗ represents pT79 plasmid; □ represents pT79GFP plasmid; ◇ represents GFPGH1-WT; ○ represents GFPGH1-HP; △ represents GFPGH1-3P).

3.3.6 Prediction of intermolecular interaction between 5′-end of GH1 mRNA and 3′-end of 16S rRNA

The results described above have demonstrated that the first 4 codons (CCC GCG GGC CCC) play a decisive role in controlling GH1 expression in *E.coli.* A number of studies have suggested that interactions between mRNA and the 16S ribosomal RNA, other than those involving the Shine-Dalgarno sequence and the initiator codon, may be involved in modulating the initiation of translation (Gold *et al.*, 1984; Gold, 1988; Makrides, 1996; Olins & Rangwala, 1989; Sprengart *et al.*, 1990; Sprengart *et al.*, 1996). For this reason, the 16S RNA sequence was searched for complementarity to the region of the GH1 coding sequence which had been subject to mutagenesis (+3 to +15). A part of the 16S molecule was identified (1526 to 1510) which, after allowing for two small bulges, showed almost perfect complementarity to the pertinent region (+5 to +19) of wild-type GH1 mRNA (Figure 3-21). This complementary region of the 16S RNA molecule lies only a few nucleotides 5′ of the anti Shine-Dalgarno sequence and comprises most of a short inverted repeat predicted to form a hairpin structure (Figures 3-21 & 3-22). The calculated minimum free energy for duplex formation between the mRNA and the 16S RNA over the region of complementarity (not including the Shine-Dalgarno region) amounts to -25.8 kcal/mol (Figure 3-23). This suggests that hybrid formation would be substantially more favourable than 16S intramolecular hairpin formation given that the minimum free energy for the latter structure is only -10.1 kcal/mol.

The above observation raises the possibility that inappropriate hybrid formation between GH1 mRNA and the 16S RNA molecule could be responsible for preventing GH1 expression. If this were the case then one would predict that mutations introduced into the GH1 coding region which reduce the predicted stability of the mRNA-16S RNA hybrid, should increase GH1 expression. As shown in Figure 3-23, with the exception of a few data points (discussed subsequently), this is indeed the case.

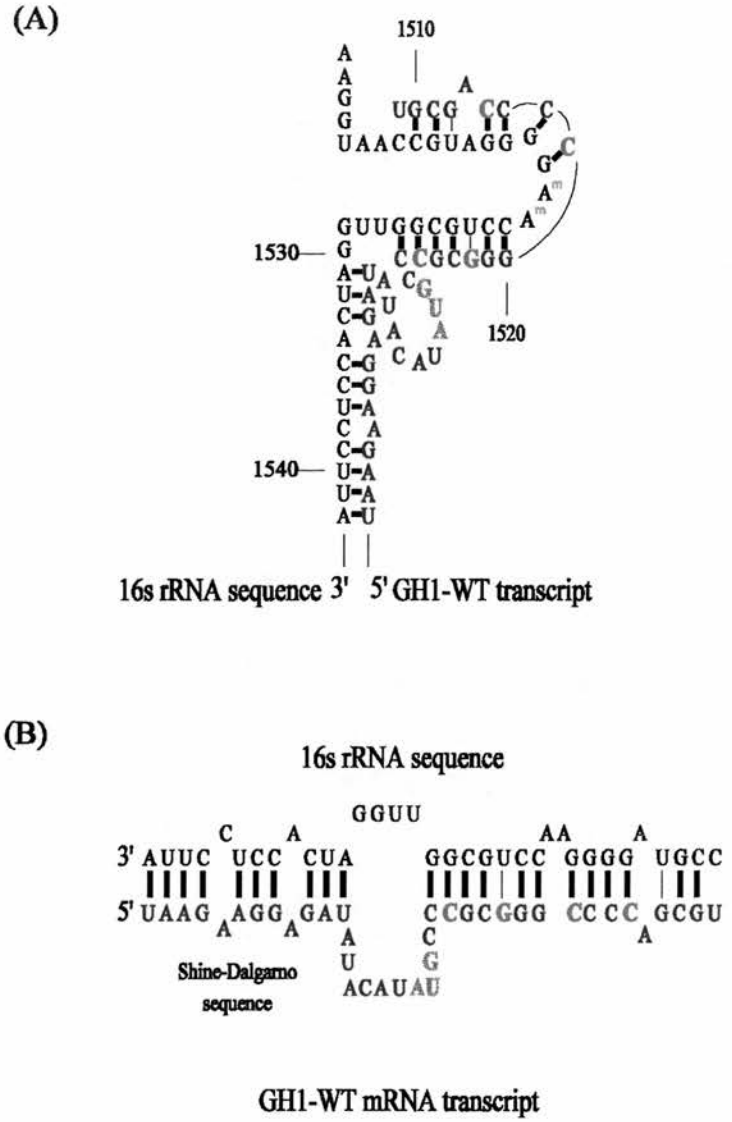


Figure 3-21 Potential hybrid formation between GH1-WT mRNA and the 3' end of *E.coli* 16S rRNA. The complementarity is shown (A) in the context of the hairpin at the 3' end of the 16S molecule and (B) in a simplified, expanded form. Colour coding has been employed to highlight the 16S sequence (black), the mRNA sequence (blue), the initiator codon (purple), and the bases in the GH1 mRNA which were mutated in the constructs(red).

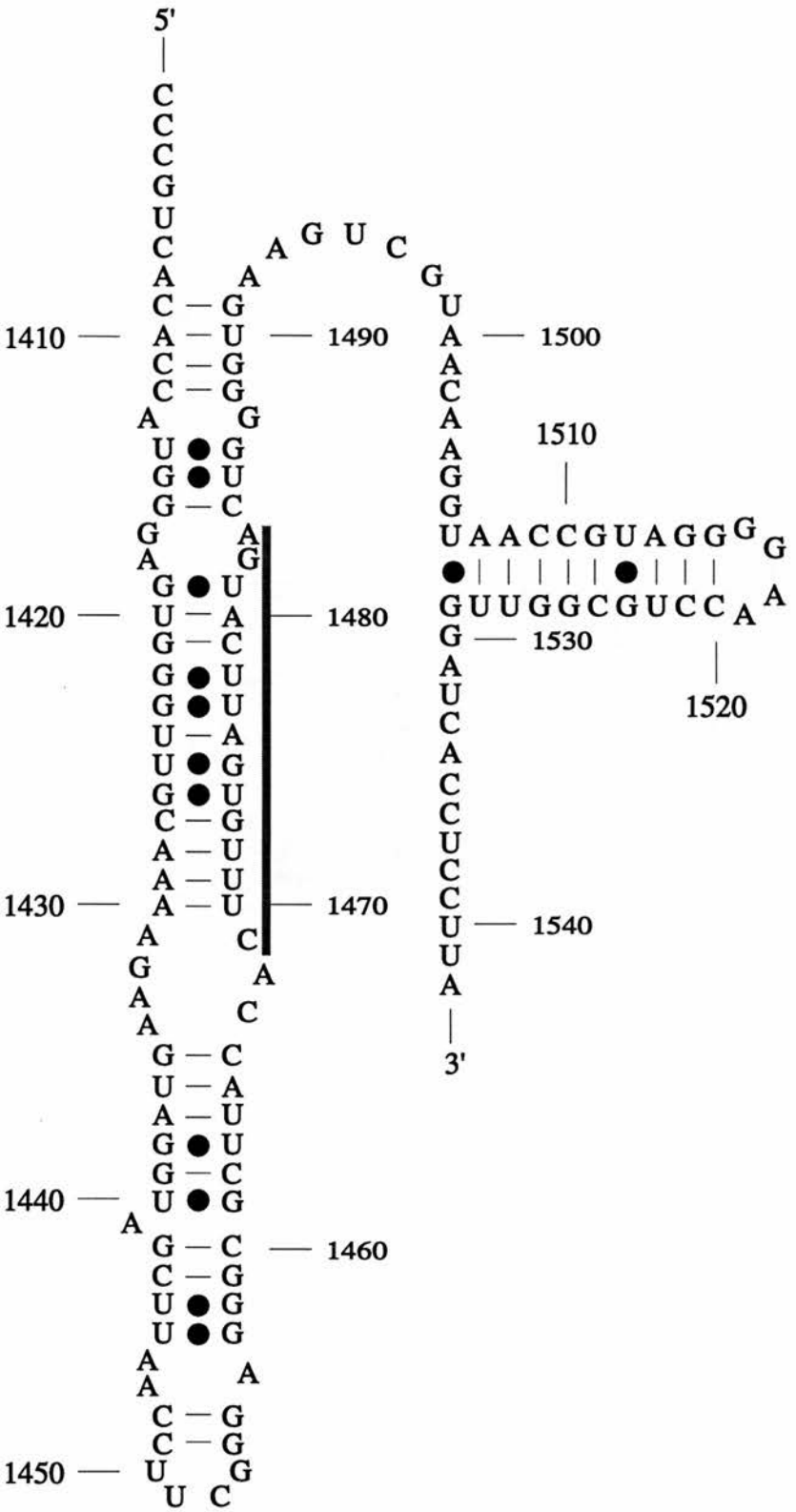


Figure 3-22 Secondary structure at the 3' end of *E.coli* 16S rRNA. The decoding region extends from nucleotide 1494 to the 3' end of the molecule. (Adapted from Woese *et al.*, 1990). The bar indicates the anti-downstream box.

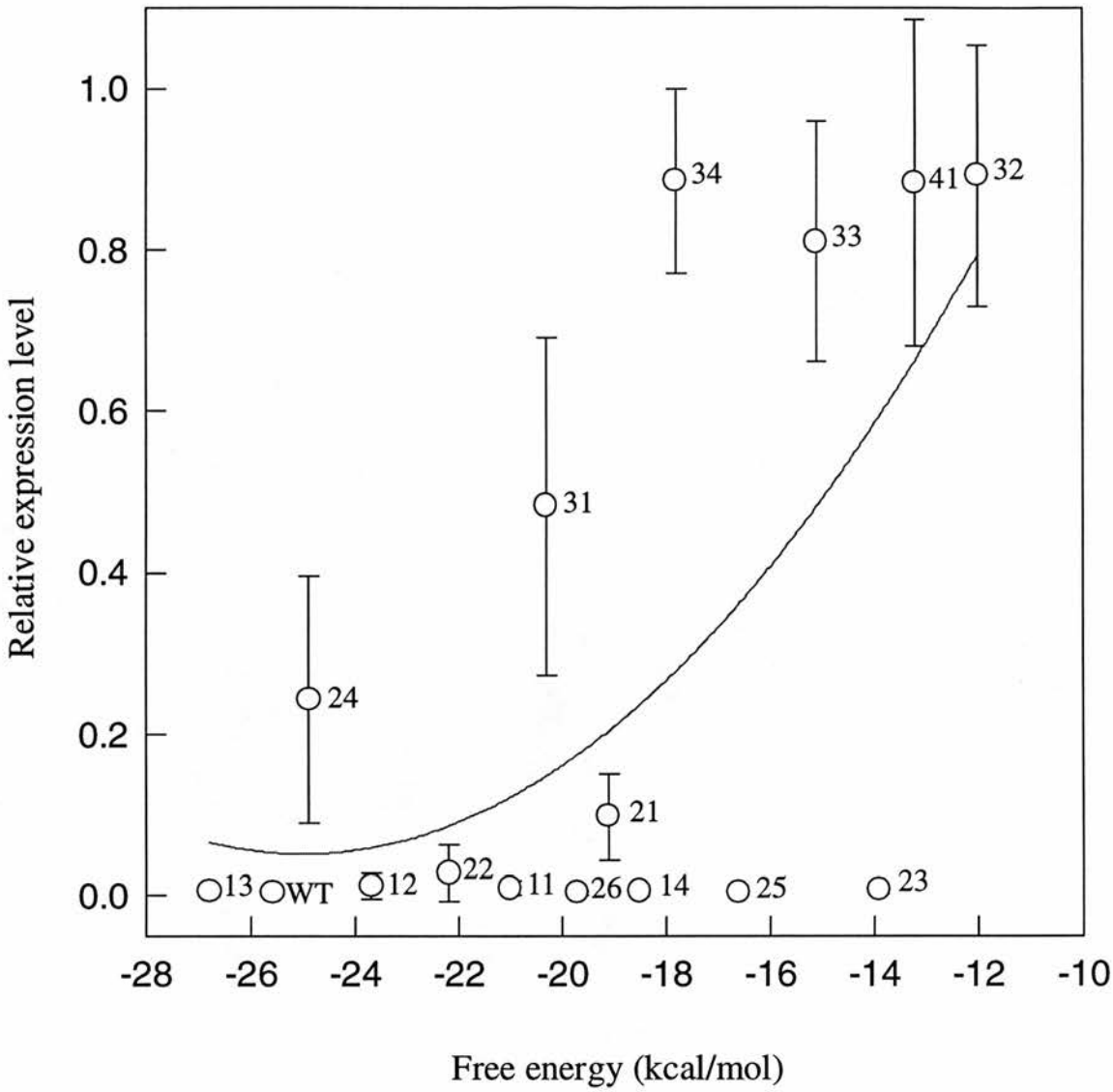


Figure 3-23 Correlation between the expression of GH1 (from Figure 3-9) and the free energy of hybrid formation between GH1 mRNA (+4 to +19) and the hairpin region of 16S rRNA (1526 to 1509). Each sample point is labelled in the right hand side.

3.4 Discussion

These studies have shown that although the wild-type coding sequence for the globular domain of the chicken linker histone H1 is very poorly expressed in *E. coli*, a considerable increase in expression can be effected by altering the first four codons of the GH1 sequence. In an attempt to explain this observation, a set of 16 GH1 expression plasmids, comprising all possible combinations of the wild-type and mutant codons, was constructed and the capacity of each variant to express GH1 was examined.

The levels of GH1 expression exhibited by the different mutants varied over a 300-fold range (Figure 3-9). This diversity could not be attributed to the abundance of GH1 mRNA (Figure 3-10). The GH1 expression level can be correlated to the AT content just downstream of the start codon, an observation consistent with previous studies (Gold *et al.*, 1981; Gerchman *et al.*, 1994). Gerchman *et al.* (1994) showed that the translation efficiency of chicken GH1 varied considerably with the nature of the codons just downstream of the ATG initiator codon, and attributed poor expression specifically to this G/C rich sequence. However, their studies did not identify what was responsible, or required for high-level expression.

Systematic site-directed mutagenesis has shed further light on this matter. The changes introduced into the GH1 coding sequence were partly chosen to overcome a potential codon usage problem. In the wild-type sequence, two of the first four codons are CCC triplets coding for proline (Table 3-1). Although this codon is rarely used in *E. coli* (Grosjean & Fiers, 1982), its occurrence in the GH1 coding sequence does not appear to be responsible for the low expression of wild-type GH1. In the mutant GH1-23, for example, both CCC codons were changed to CCA, a preferred proline codon, but the changes had no positive effect upon GH1 expression (Figure 3-9). On the other hand, In the first four codons of the wild-type sequence, GCG and GGC are the most frequently used for alanine and glycine in *E.coli*, respectively. In the mutant GH1-24, both GCG and GGC were changed to GCA and GGT, respectively. Although both

frequently used codons changed to rarely used codons, the changes had led to a substantial expression of GH1 (Table 3-1, Figure 3-9).

There is good evidence that mRNA intramolecular secondary structure is a key factor in determining the efficiency of translational initiation in prokayotes (Stormo, 1986; Gold, 1988; de Smit & van Duin, 1990b). The expression of several genes in *E.coli* appeared inversely related to the stability of the secondary structure of their ribosome binding site (Hall *et al.*, 1982; Tessier *et al.*, 1984; Buell *et al.*, 1985; Spanjaard *et al.*, 1989). Quantitative analysis of the relationship between translational efficiency and the mRNA secondary structure in the initiation region indicated that initiation is completely dependent on spontaneous unfolding of the entire initiation region, and nucleotides outside the start codon and the SD region only affect expression if their mutation changes the stability of the secondary structure (de Smit & van Duin, 1990a; de Smit & van Duin, 1990b; de Smit & van Duin, 1994a). However, no evidence was found to support that any of the 16 mRNAs generated in this study had the potential to adopt secondary structure which would compromise the interaction of the Shine-Dalgarno sequence with its complement in 16S rRNA (Figure 3-11 & 3-12). Although predicted secondary structure within the coding region of the mRNA varied with changes in sequence, the stability of these structures did not correlate to GH1 expression (Figure 3-12).

On comparing the wild-type GH1 coding sequence around the region subject to mutagenesis (+3 to +15) with the 16S ribosomal RNA sequence, a stretch of notable complementarity was identified. The section of the 16S RNA molecule concerned (1526 to 1510) is located about 7 nucleotides 5′ of the anti-Shine-Dalgarno sequence and comprises most of an inverted repeat predicted to form a hairpin structure (Figures 3-21 & 3-22). Energy calculations indicate that a hybrid formed between the GH1 mRNA and 16S RNA hairpin (Figure 3-23) would be much more stable than the 16S RNA hairpin itself (-25.6 kcal/mol compared to -10.1 kcal/mol). There are, however, likely to be other factors of the organisation of 16S RNA in the 30S subunit which could influence the hairpin stability – for example it might be buried or complexed with a ribosomal protein. In this context it is also worth considering that hybrid

formation between the Shine-Dalgarno and anti-Shine-Dalgarno sequences would bring the 16S RNA hairpin and the GH1 mRNA "anti-hairpin" sequences into very close proximity, which could enhance the likelihood of their forming this additional hybrid (Gold *et al.*, 1984; Gold, 1988). Furthermore, mutations in the GH1 coding sequence which would reduce the stability of a hybrid formed at the 16S hairpin tend to enhance GH1 expression (Figure 3-23), although the distribution of the data points in this analysis indicate that the relationship is not a simple one. Nevertheless, these observations broadly support the possibility that hybrid formation between GH1 mRNA and the hairpin sequence of the 16S RNA molecule could be involved in the failure to express GH1 protein from the wild-type sequence. Two of the experiments undertaken are consistent with this model. Firstly, the analysis of mRNA distribution within ribosomal subunits showed that for the low-expressing plasmid (GH1-WT), most of the GH1-WT mRNA was found in the 30S subunit whereas for the high-expressing plasmid (GH1-32), very little of the mRNA was found in 30S subunit (Figures 3-14, 3-15 & 3-16). This result indicated that GH1-WT mRNA was prevented from transferring into a 70S subunit. Secondly, results from the introduction of the first four codons of GH1 in front of the GFP gene indicate that the diversity of expression levels of GFP fusion constructs is inherent from these 12 nucleotides (Figures 3-18, 3-19 & 3-20). For example, GFPGH1-WT could not express at all, whereas GFPGH1-3P could express very well (Figures 3-18, 3-19 & 3-20). Both of these results are consistent with the proposal that there may be an interaction between the 16S rRNA and first four codons of GH1 which causes a detrimental effect on expression.

The above proposition raises two questions. Firstly, how could the mRNA-16S RNA hybrid come to be formed, given that the hairpin might well be an established feature of the 16S RNA in the 30S ribosomal subunit? Secondly, if a hybrid does form, why should it have such a pronounced influence upon translation?

Although the data demonstrate an association between the stability of the potential hybrid and the level of GH1 expression (Figure 3-23), there is a problem in attributing expression solely to duplex formation. Given that the stability of the hybrid,

irrespective of mutations (Figure 3-23), is always greater than that of the hairpin (-10.1 kcal/mol), a hybrid might be expected to form in all cases and GH1 should not be expressed in any instance, which is clearly not the case. Therefore, hybrid formation cannot be driven by an equilibrium process governed simply by the relative minimum free energy values of the hybrid and hairpin structures.

The formation of the postulated mRNA-16S RNA hybrid might be initiated (nucleated) within the loop of the 16S RNA hairpin (Figure 3-22, 1516-1519 nt). Indeed, there are aspects of the results which suggest that this may be a critical step in the pathway to hybrid formation. None of the individual mutations or pairs of mutations which reduce complementarity exclusively within the stem region of the 16S hairpin have any substantial impact upon GH1 expression (GH1-11, 13, 14, 22, 23 or 26, Table 3-1, Figures 3-9 & 3-21A). GH1-23 is particularly notable in this context as it reduces the minimum free energy of the proposed hybrid from -25.6 to -13.9 kcal/mol.

A simplistic pathway for the formation of an mRNA-16S RNA hybrid, involving nucleation within the 16S hairpin loop, is outlined in Figure 3-24. In one scheme the reactive mRNA is considered to be in a relatively unstructured form which should present no particular barrier to the initiation of hybrid formation. This will involve base-pairing between the mRNA dinucleotide $C/U_{12}C_{13}$ (depending on the mutant in question) (Figure 3-21 A) and the pair of G residues in the 16S hairpin loop. Hybridisation can then proceed by zippering of the mRNA and rRNA down the stem of the hairpin. Under these conditions, hybrid formation and GH1 expression should be a function of the relative minimum free energy values of the hybrid and hairpin structures. As argued above, the data do not support this interpretation.

In the alternative pathway (Figure 3-24), the reactive mRNA is considered to be in secondary structure, which is likely to be the case. Here the capacity to initiate hybrid formation with 16S RNA will depend upon the availability of the $C/U_{12}C_{13}$ dinucleotide in the mRNA. If it is contained within a loop, hybrid formation may proceed, whereas if it is contained within a stem structure, hybrid formation may not be

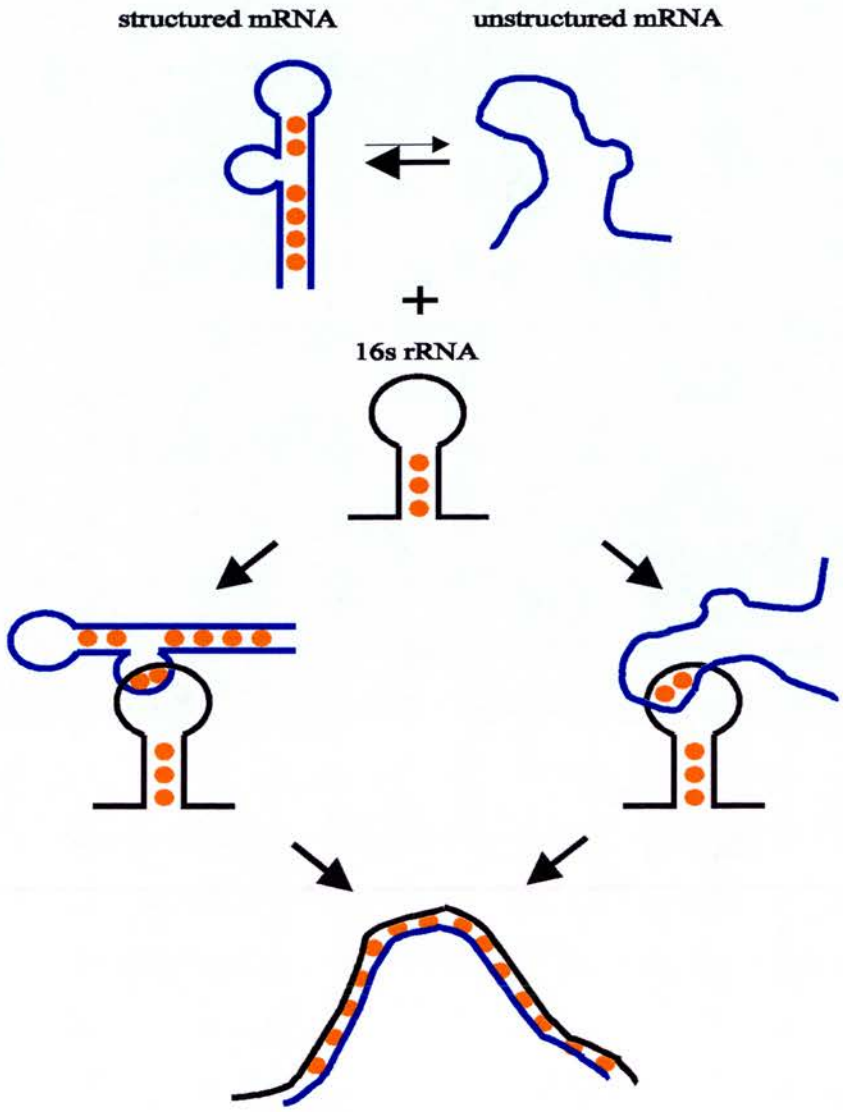


Figure 3-24 Proposed formation of hybrid between 16S RNA and mRNA. The 16S RNA is black and mRNA is blue. Base pairing is indicated by orange filled circles.

possible. Sequence changes have the potential to alter mRNA secondary structure and by this means could influence hybrid formation and consequently, GH1 expression.

The secondary structures predicted by MFOLD for all but one of the GH1 mRNAs (-43 to +47) fall into three forms (Figure 3-25). GH1-33 occupies a distinct class (Figure 3-11), but for the purpose of this discussion can be considered as Type II. With respect to the location of the $C/U_{12}C_{13}$ dinucleotide, the three forms of secondary structure are quite distinct. In Type I, the dinucleotide is located centrally within a single-stranded region of a bulge structure and, in the context of the above argument, could therefore be available to initiate hybrid formation. None of the seven GH1 sequence variants which adopt the Type I structure express GH1 (Figure 3-9). In the Type II structure, the $C/U_{12}C_{13}$ dinucleotide is entirely contained within a stem region and should not be available to initiate hybrid formation. All 5 of the mutants which fall into this category express GH1 (Figure 3-9). In Type III, the nucleating dinucleotide is partly in a stem and partly in a bulge. There are 4 members of the Type III class, two of which express GH1 and two which do not (Figure 3-9).

The above analysis has been repeated using a shorter section of the GH1 mRNA (-43 to +37) for secondary structure prediction. Even though a greater variety of structures were predicted, the relationship between availability of the $C/U_{12}C_{13}$ dinucleotide and expression of GH1 was upheld.

If hybrid formation between GH1 mRNA and 16S rRNA is partly controlled by the secondary structure of the former component, this feature of the pathway could explain some of the apparently aberrant data points in Figure 3-23. For example, GH1-23, as a member of the Type I secondary structure class (Figure 3-25), has the potential to initiate hybrid formation which, despite having a relatively high minimum free energy, must be sufficient to establish duplex formation and prevent GH1 expression. Similar arguments can be applied to explain the properties of GH1-24, GH1-25 and GH1-31.

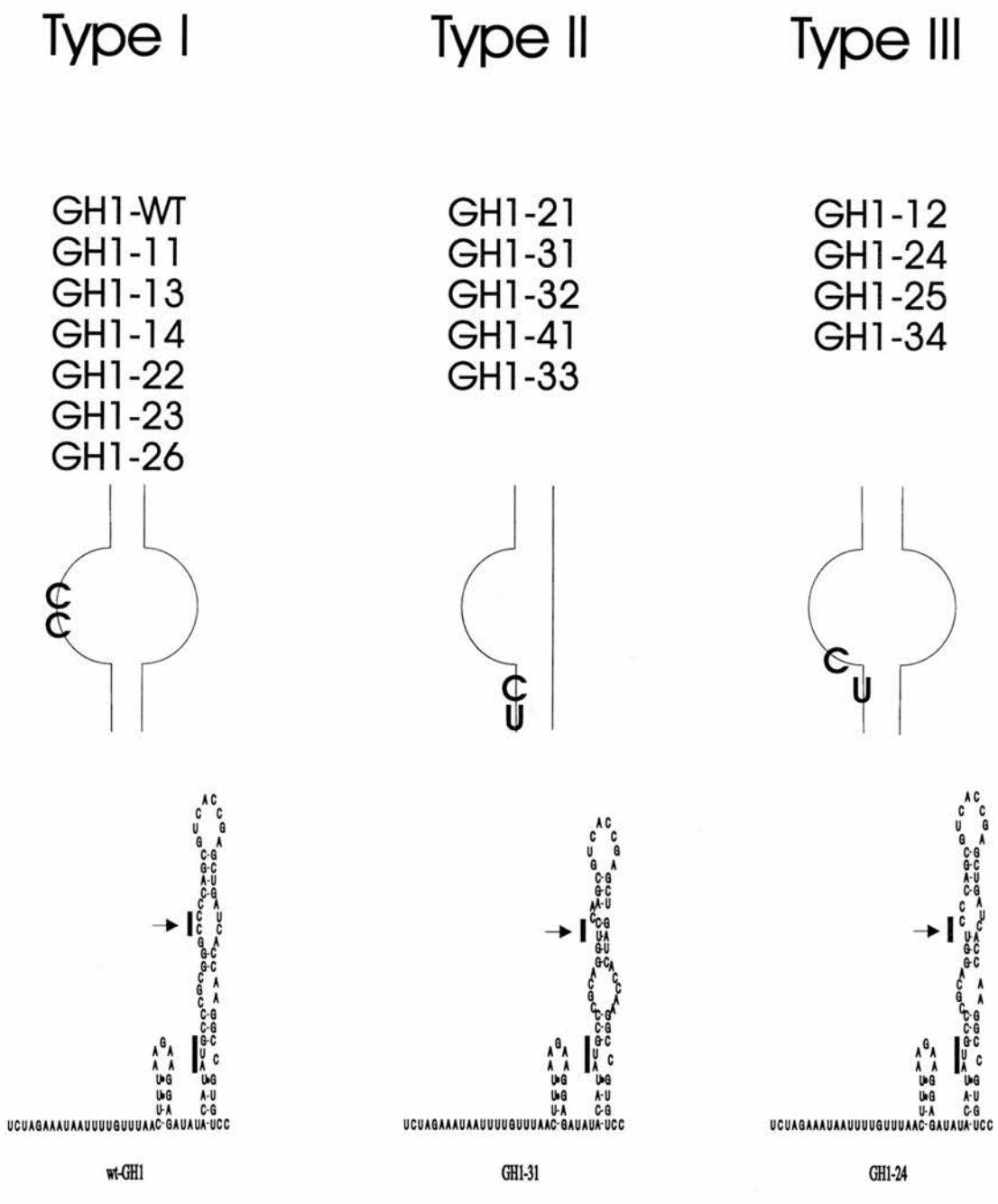


Figure 3-25 Classification of predicted mRNA secondary structures in three classes according to CU/C dinucleotide exposure.

In the GFPGH1 fusion expression experiments the oligonucleotide GH1-HP (Figure 3-5) was employed to test the proposal that the capacity to initiate hybrid formation with 16S RNA depends upon interaction with the GG sequence in the 3′-end hairpin of 16S RNA. The mutation introduced into the GH1 oligo GFPGH1-HP ($C_{12}C_{13}$ to $G_{12}G_{13}$) should eliminate any possibility of hybrid formation at this location. The expectation would be that the mutant should express well. However, this fusion plasmid did not express at all (Figures 3-18, 3-19 & 3-20). These results may suggest that the capacity to initiate hybrid formation with 16S rRNA does not merely depend upon the availability of the $C/U_{12}C_{13}$ dinucleotide in the mRNA.

Numerous previous studies have suggested that interactions between mRNA and 16S RNA, other than at the Shine-Dalgarno site, might contribute to the regulation of translational initiation (Gold *et al.*, 1984; Gold, 1988; Olins & Rangwala, 1989; McCarthy & Gualerzi, 1990; Sprengart *et al.*, 1990; Sprengart *et al.*, 1996). In many of these cases, the proposal has been based on observations of complementarity provided by computer analysis of sequence data. Invariably, the consequence of interaction between mRNA and 16S RNA is proposed to confer an enhancing effect upon initiation although details of the possible mechanism involved remain unclear. One notable exception is the scheme proposed by Gold *et al.* (Gold *et al.*, 1984; Gold, 1988), to explain the unusual sequence of *E. coli* IF3 mRNA. The model invokes extensive annealing between mRNA and 16S RNA including an instance of hybrid formation which appears to be initiated within a 16S hairpin loop (residues 462-470; Figure 3-22) and involves at least partial disruption of the hairpin stem (Gold *et al.*, 1984). This, and the other schemes which invoke mRNA interaction with 16S RNA, are founded on the idea that regions of the 16S RNA molecule are exposed on the surface of the ribosome and that when mRNA becomes attached during initiation some of its sequence will be proximal to those exposed 16S nucleotides thus presenting an opportunity for hybrid formation. In bacterial mRNAs, such a potential could be utilised advantageously (Gold, 1988). On the other hand, mRNA interactions with 16S RNA which would be detrimental to initiation or translation would have been selected against. However, when foreign mRNAs, such as those derived from eukaryotic genes, are introduced into *E. coli*, the opportunity for inappropriate hybrid formation which

has a deleterious effect on translation, must become a distinct possibility. Observations with GH1 mRNA in GH1 site-directed mutagenesis experiments (Figure 3-9), GFPGH1 fusion protein expression experiments (Figures 3-18, 3-19 & 3-20), and ribosomal RNA fraction results (Figures 3-13, 3-14, 3-15 & 3-16) may reflect such a case.

In both prokaryotes and eukaryotes, mRNAs are recruited to ribosomes in a sequential, multistep process. In eukaryotes, following the recruitment of the small ribosomal subunit to the mRNA, the mRNA sequence is scanned and the small subunit is placed at the initiation codon. After this, the joining of the large ribosomal subunit to the mRNA completes the assembly of the ribosome. According to the scanning model of eukaryotic translation (Kozak, 1989), the 40S ribosomal subunit with its associated factors engages the mRNA at or near the methyl guanosine cap and then scans in a 3′ direction. Upon encountering the first initiation codon, the 60S subunit joins the 40S subunit to form a complete 80S ribosome, and polypeptide synthesis commences. Therefore, there are substantial differences between prokaryotes and eukaryotes in the initiation of translation: in prokaryotes the control is primarily directed by the base-pairing between the 16S rRNA and the Shine-Dalgarno or downstream coding sequence on mRNA; in eukaryotes the control is primarily directed by protein-protein and protein-RNA interactions (Lamphear *et al.*, 1995; Merrick & Hershey, 1996; Hentze, 1997). For all cellular and the vast majority of viral mRNAs, it thus seems unlikely that initiation site recognition is by base-pairing between 18S rRNA elements and one or more mRNA motif, although the 3′ end of the 18S subunit is conserved in both eukaryotes and prokaryotes. Nevertheless, a possible exception is clearly those RNAs translated by internal initiation, where a *cis*-acting RNA element , which is generally known as the IRES (Internal Ribosome Entry Site), can promote direct ribosome entry to the initiation site of the downstream cistron in di- or even tricistronic mRNAs. It is likely that both RNA-RNA and protein-RNA complexes are involved in this translation initiation (Chen & Sarnow, 1995). In addition, both the pyrimidine-rich sequence element and sequences surrounding the start-site AUG codon have been predicted to be complementary to 18S rRNA (Pilipenko *et al.*, 1992). However, an interaction between the 18S rRNA and mRNA remains to be confirmed,

although the sequence 980-1061 in the central domain of human 18S rRNA has been proved to have an inhibitory action upon translation initiation when sequestered with complementary DNA probes (Graifer *et al.*, 1997).

The prediction of an intramolecular interaction between GH1 mRNA and the 16S RNA molecule has raised the possibility that aberrant hybrid formation may control translation initiation in this case. The sequence of 16S rRNA 1510-1526 of *E.coli* is highly conserved in eukaryotic 18S rRNA (Gutell *et al.*, 1985). Therefore, it was worth testing whether the same interaction would form in an eukaryotic expression system. For this *in vitro* eukaryotic transcription/translation experiment, the GFPGH1 fusions and GH1 mutants GH1-WT, GH1-3P, GFPGH1-WT and GFPGH1-3P DNAs were used as template separately in TNT T7 Quick Coupled Transcription/Translation System (Promega). The results showed that all these GH1 constructs were expressed to the same level (Figure 3-26).

This result indicates that translation initiation in an eukaryotic system is not influenced by the sequence immediately downstream of the start codon. In an eukaryotic expression system the main factors recognised by 18S rRNA are the m7G cap and the 5′ -untranslated region of the mRNA. Therefore, it is quite reasonable that all plasmids expressed to the same level.

It could be suggested that bacterial mRNAs would avoid using 5′ coding sequences which complement the 16S RNA hairpin. The fact that the first 20 or so nucleotides following the initiation codon of prokaryotic mRNAs display a very distinct bias for AT richness (Schneider *et al.*, 1986) is consistent with this idea given that the 16S RNA hairpin is particularly GC rich (Figure 3-21). Conversely, in the case of foreign genes introduced into *E. coli*, it is well established that a GC-rich initial coding sequence can often prove to be problematic in terms of protein expression. The aberrant hybrid proposal could clearly have some relevance for this latter phenomenon.

If the formation of a hybrid between GH1 mRNA and the 16S hairpin is responsible for the failure of GH1 mRNA to be translated, it could effect this action in

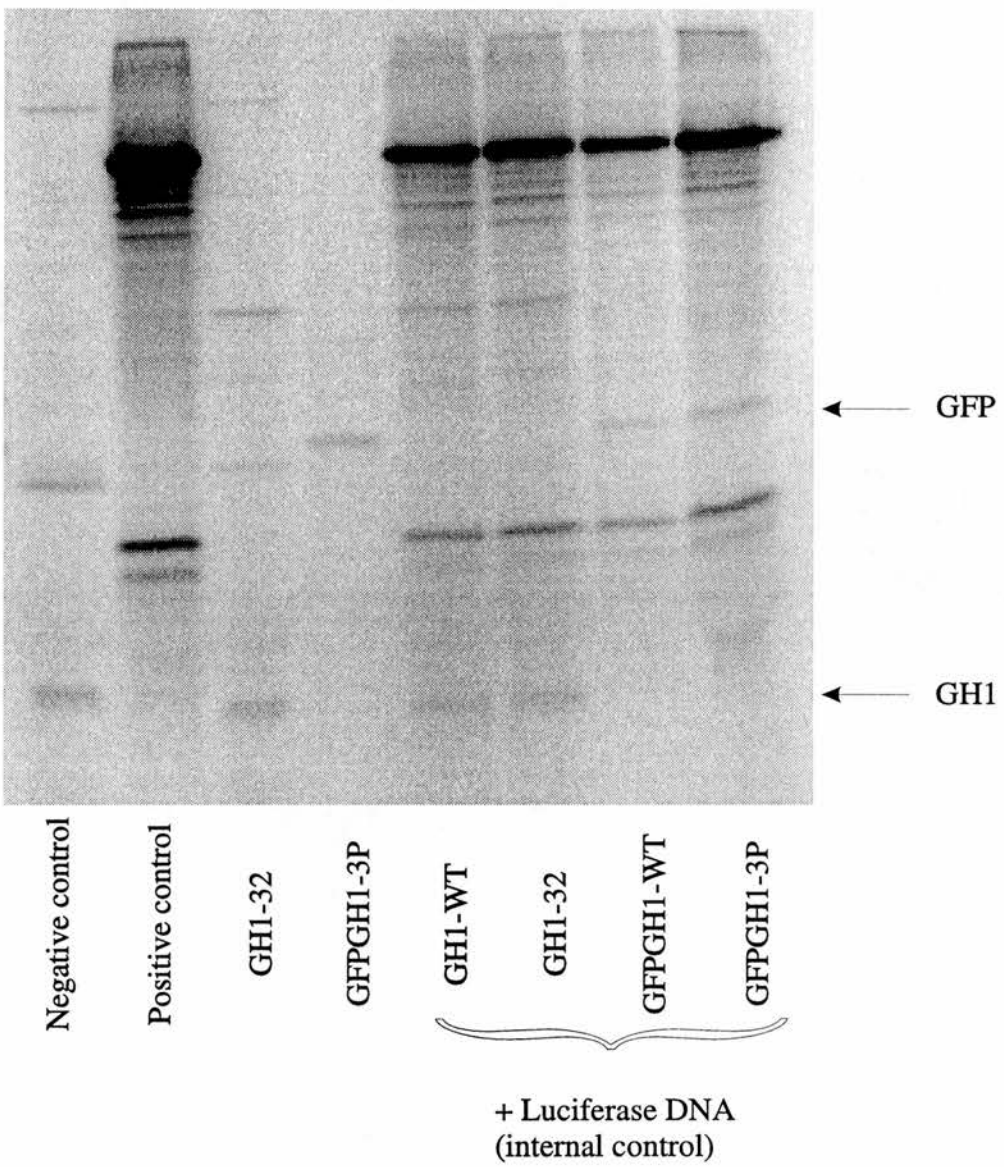


Figure 3-26 *In vitro* transcription/translation of GFPGH1 fusion and GH1 constructs under eukayotic transcription/translation system. The Luciferase DNA was added in the positive control experiment whereas no any DNA was added in the negative control experiment.

a number of ways. Firstly, the stability of the hybrid could simply prevent the process of translational elongation. Secondly, fMet-tRNA^{fMet} access to the initiator codon in the mRNA could be blocked. Alternatively, disruption of the 16S hairpin may fundamentally compromise the process of translational initiation. Support for the latter proposal can be inferred from the fact that the 48 nucleotide region at the 3′ end of the 16S RNA molecule, often referred to as the "decoding" region (Gold, 1988; McCarthy & Gualerzi, 1990; Makrides, 1996) (Figure 3-22), is known to adopt distinct structural configurations when the 30S ribosomal subunit is in an active as opposed to an inactive state (Weller & Hill, 1992). The 3′ hairpin, focal to the interpretation of these results, is located in the centre of this dynamic, decoding region. The model proposed to explain the translational properties of GH1 mRNA is detailed in many respects and as such lends itself to further experimental testing, for example extension inhibition experiments (toeprinting) could identify this kind of interaction between mRNA and 16S RNA.

Chapter 4 Effects of Linker Histones on Nucleosome Positioning over the Chicken β-globin Gene

4.1 Introduction

Most nucleosomes adopt a unique stable position with respect to their underlying DNA sequence. It follows that DNA sequences possess the inherent ability to position nucleosomes precisely. (Rhode, 1985; Bergman, 1986; Richard-Foy & Hager, 1987; Hayes *et al.*, 1990; Schild *et al.*, 1993; Jackson & Benyajati, 1993). The precise positioning of the histone octamer on eukaryotic DNA is a crucial element in the regulation of eukaryotic genes (Felsenfeld, 1992). For example, positioned nucleosomes can influence the transcription of chromatin templates by virtue of their ability to prevent the access of *trans*-acting factors to DNA (Simpson, 1991; Lee *et al.*, 1993; Li & Wrange, 1993). Many explanations have been proposed to describe the change in the contacts between DNA and histone octamer that occur during gene activation (Thoma, 1991; Hayes & Wolffe, 1992b; Adams & Workman, 1993; Chen *et al.*, 1994; Wall *et al.*, 1995; Walter *et al.*, 1995; Studitsky *et al.*, 1995), but the structural features of the nucleosome that might contribute to transcriptional regulation are not clear yet. To determine whether there are features in the long-range organisation of DNA sequence which could influence the higher-order packaging of chromatin and thereby contribute to the regulation of gene expression, it is, therefore, very important to produce a long-range, high-resolution nucleosome positioning map for an entire gene region.

4.1.1 Definition of nucleosome positioning

The location of DNA on nucleosomes or the positioning of a histone octamer on a particular DNA sequence can be described in two ways – translational positioning and rotational positioning (Figure 4-1). Translational positioning and rotational positioning are inherently related. The description of the position of DNA with regard to the boundaries of the nucleosome is called translational positioning

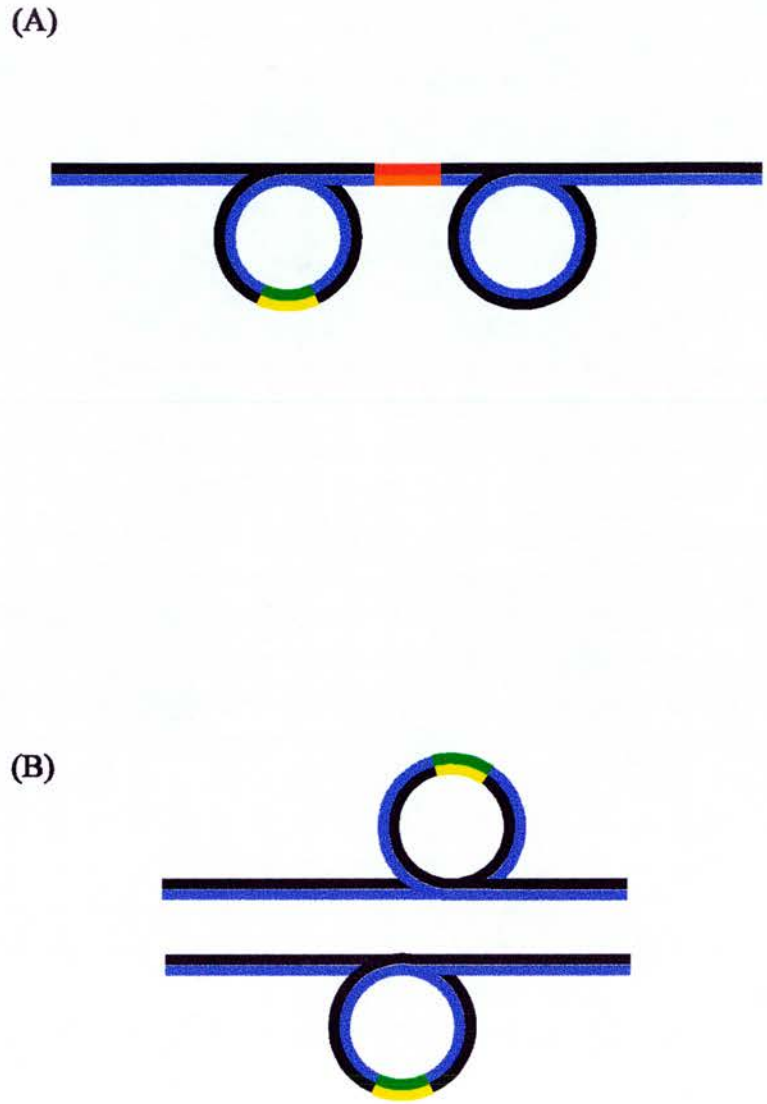


Figure 4-1 Illustration of nucleosome positioning parameters. (A) Example of translational positioning with respect to a particular region on the linker DNA. (B) Example of rotational positioning with respect to a particular segment of nucleosomal DNA which can face inward or outward from the histone octamer.

(Travers & Klug, 1987). The orientation of the double helix with regard to the octamer surface is called rotational positioning (Travers & Klug, 1987). Rotational positioning is very important because it defines the side of the helix that is bound to the protein surface and is, therefore, inaccessible, whereas the side that is exposed is more accessible. Figure 4-1A shows two nucleosome positioning sites relative to one specific region. Figure 4-1B shows the orientation of a specific DNA region to the histones. If the core particle is moved 5 bp, the contact between the DNA and the histone octamer would change maximally and therefore the accessibility of this DNA region is changed. This is the so-called rotational setting.

Many studies have suggested that dynamic features of nucleosome positioning can contribute to the regulation of gene expression (Pennings *et al.*, 1991; Wall *et al.*, 1995; Varga – Weisz *et al.*, 1995). As the nucleosome may dominate the transcription by virtue of their sliding and transient dissociation within the transcription factor binding region, it is important to define nucleosome positioning with respect to the underlying DNA

4.1.2 Detection of nucleosome position

A combination of micrococcal nuclease (MNase) digestion and indirect end-labelling methods are usually employed to determine the location of nucleosomes *in vivo* (Wu, 1980; Thoma *et al.*, 1984). In chromatin, MNase cuts preferentially in the linker DNA between nucleosomes. Therefore, when the chromatin is first digested with MNase and subsequently cut with a restriction endonuclease this allow nucleosome binding sites or the linker between the nucleosomes to be located. Apart from MNase, in combination with a restriction endonuclease (Almer *et al.*, 1986; Reik *et al.*, 1991), DNase I or a chemical cleavage reagent (Cartwright *et al.*, 1983) can be used to determine nucleosome positions. However, both nucleosome mapping methods are relatively low resolution. In the study of nucleosome positioning in a yeast minichromosome, Shimizu *et al.* (1991) have developed a primer extension assay which allows a much more detailed examination of the

location of nucleosomes. Their work indicates that nucleosomes are precisely and stably positioned both translationally and rotationally over the sequences adjoining the operator of the yeast α-2 repressor gene.

For *in vitro* mapping, MNase, DNase I or exonuclease III have been used for nucleosome mapping at high resolution. These analyses are usually applied to short, end-labelled linear DNA fragments which have been reconstituted with purified core histones (Drew & Travers, 1985; Drew & Calladine, 1987; Kefalas *et al.*, 1988; Hayes *et al.*, 1990; Buckle *et al.*, 1991). Recently, a base-pair resolution method for determing nucleosome positions *in vitro* has been developed. The procedure is based on the ability of a ferrous-ion-chelating reagents, tethered to DNA-binding proteins, to cut DNA in the vicinity of the modification site through generation of hydroxyl radicals (Flaus *et al.*, 1996).

Recently, the monomer extension method has been developed to produce a long-range nucleosome positioning map for the entire chicken β-globin gene (Yenidunya *et al.*, 1994; Davey *et al.*, 1995). The technique maps the precise translational positions adopted by core histone octamers reconstituted onto long DNA sequences and maps the boundaries of core particle DNA fragments protected by the histone octamer from micrococcal nuclease digestion. This technique can be used to reveal binding sites over extensive stretches of a single reconstitute and provide high resolution *in vitro* mapping data.

Much evidence has been gathered to show that nucleosomes can adopt well-defined locations with respect to DNA sequence both in reconstituted systems and within the nucleus (Bock *et al.*, 1984; Thoma & Simpson, 1985; Ramsay, 1986; Drew & Calladine, 1987; Perlmann & Wrange, 1988; Pina *et al.*, 1990; Buckle *et al.*, 1991). The determination of *in vitro* nucleosome positioning is an important approach, because it is frequently if not invariably found that positions mapped *in vitro* correspond to those which can be determined *in vivo*. Therefore, information obtained from *in vitro* mapping should reflect the functional significance of nucleosome positioning *in vivo*.

4.1.3 Mechanisms of nucleosome positioning

A large fraction of the genome might be organised into positioned nucleosomes (Pratt & Hattman, 1983). Since all cells have the same DNA sequence but can adopt different repeat lengths/positions, the contribution of positioning information in the DNA sequence must be modulated by other factors *in vivo*. Many factors have been proposed to influence the formation of nucleosome positioning. These factors include DNA sequence or structure, non-histone proteins and linker histones. Therefore, the properties of DNA itself are not sufficient for nucleosome positioning, but also need other factors to help determining the position.

4.1.3.1 Boundary-directed nucleosome positioning

It has been proposed that nucleosome positioning can be directed by boundaries. There is evidence to support the fact that the binding of DNA sequence-specific non-histone proteins can act as a boundary element. For example, an abundant yeast protein GRF2, binds to the upstream activating sequence, acts as an auxiliary gene activator, and strongly dictates adjacent nucleosome positioning (Fedor *et al.*, 1988; Chasman *et al.*, 1990). The binding of α2, a yeast repressor, causes nucleosomes to be precisely positioned downstream of the operator sequences (Roth *et al.*, 1990; Shimizu *et al.*, 1991).

Kornberg & Stryer (1988) have proposed that the binding of a sequence-specific protein to DNA creates a boundary whose effect upon neighbouring nucleosomes is of first-order near the boundary and decays to lower order with increasing distance from the boundary. The preferential binding of histones to certain sequences is a second-order effect, whose influence upon neighbouring nucleosomes is then of third-order.

It has also been argued that DNA structure may act as boundaries. For example, cruciform DNA, Z-DNA formed on poly (dG•m^5dC)• poly (dG•m^5dC) or long stretches of poly (dA)• poly(dT) which are unable to associate with histones *in*

vitro (Nickol *et al.*, 1982; Nobile *et al.*, 1986), might act as boundaries to position or exclude nucleosomes.

Although the details of the higher order chromatin organisation are not clear, the folding of the nucleosomes into a higher order compact structure might affect the nearest neighbour nucleosome interactions, and then modulate nucleosome positioning (Thoma *et al.*, 1979; Thoma & Koller, 1981; Thoma & Zatchej, 1988). In addition, nuclease-sensitive regions (Thoma, 1986; Perez-Ortin *et al.*, 1989; Bernardi *et al.*, 1992) and linker histones, which may dictate the formation of nucleosomes in higher order structure (Satchwell *et al.*, 1986; Thoma & Zatchej, 1988), could also be regarded as boundaries to affect nucleosome positioning.

4.1.3.2 DNA sequences as a direct determinant of nucleosome positioning

Many studies have shown that nucleosome positioning can be effected directly by DNA sequence or structure (Drew & Travers, 1985; Stachwell *et al.*, 1986; Calladine & Drew, 1986; Travers & Klug, 1987; Simpson, 1991). The nucleosomes in chromatin are associated with many different DNA sequences and the sequence-dependent preference for location of nucleosomes on DNA is likely to contribute to the control of gene regulation.

DNA bendability or the DNA bending propensity and DNA deformability are the most frequently discussed factors through which DNA sequence influences nucleosome positioning. Double-stranded DNA is usually described as a wormlike chain model with a persistence length of about 500 Å in 0.1 M NaCl (Hagerman, 1988). However, the DNA in a nucleosome is uniformly bent in a circular trajectory with a radius of 44 Å. Therefore, the DNA in a nucleosome is tightly bent in comparison to its persistence length and this folding must be energy demanding. Trifonov (1980) were the first to propose that a major determinant of nucleosome positioning was the bendability or anisotropic flexibility of DNA. The concept of bending is that certain base sequences are associated with an intrinsic curvature that leads to bending. DNA bending potential, depending on the fluctuations of the base

sequence, has been used to predict virtual nucleosome positioning from the nucleotide sequence (Turnell & Travers, 1992; De Santis *et al.*, 1993).

Intensive studies have shown that nucleosomal rotational positioning often depends on DNA anisotropic bendability (Travers & Klug, 1987; Travers, 1989). The anisotropic flexibility of a DNA molecule is a sequence-dependent property that is determined by the physicochemical characteristics of individual base steps in the sequence. The identification of the sequences responsible for the anisotropic flexibility and their importance in the rotational setting of the DNA around the histone octamer has been well established (Satchwell *et al.*, 1986). The concept of anisotropic bending is that the stiffness of a DNA fragment depends on the cross-section with respect to the axis of bending. Thus, if the DNA fragment is free to change its attitude by rotating about its axis, it adopts a configuration in which it bends about its weak axis (Calladine & Drew, 1986). A special case of anisotropic bending is the instance of inflexibly bent DNA. Curved DNA is found to fold more easily onto the histone octamer than non-curved molecules (Pennings *et al.*, 1989; Costanzo *et al.*, 1990). The features responsible for intrinsic DNA curvature have been used as a basis for the prediction of nucleosome positioning (De Santis *et al.*, 1993). These and others studies demonstrate that the structure and arrangement of DNA sequence provides information for nucleosome positioning, and confirm the relationship between rotational setting of the histone octamer and the sequence-dependent anisotropic flexibility of nucleosomal DNA (Drew & Traver, 1985; Stachwell *et al.*, 1986; Neubauer *et al.*, 1986; Drew & Calladine, 1987; Kefalas *et al.*, 1988).

4.1.3.2.1 DNA curvature: static state studies

Sequence-dependent bending in free DNA can be described by two basic models which explain the intrinsic curvature of the double-helix. The Junction model relates the curvature to deflections at each junction between the axes of two segments of B form DNA, which means that macroscopic curvature is a result of bending between straight DNA segments (Wu & Crothers, 1984; Koo *et al.*, 1986; Koo & Crothers, 1988). The other model, Nearest neighbour, proposes that smooth

bending along the double helix is caused by small additive wedge. In this proposal, macroscopic curvature of a DNA segment arises as a result of small changes in helical twist, roll, tilt and slide parameters that are considered inherent properties of dinucleotides (Trifonov, 1980; Calladine *et al.*, 1988; Cacchione *et al.*, 1989; De Santis *et al.*, 1990; Bolshoy *et al.*, 1991). Theoretically, with a proper choice of the bending parameters at given dinucleotide steps, both models can describe the same macroscopic curvature. Both models produce a net curve toward the minor groove at the centre of an A-tract, even though the underlying predicted microscopic bends may be different for each model (Crothers *et al.*, 1990; Goodsell & Dickerson, 1994). As the junction model mainly attributes the DNA curvature to periodic A tracts, DNA molecules without periodic A tracts are predicted to lack intrinsic curvature. However, some evidence has shown that molecules lacking A tracts can be curved (Diekmann, 1987; Diekmann & McLaughlin 1988; McNamara *et al.*, 1990; Bolshoy *et al.*, 1991). In addition, the junction model fails to accommodate the fact that the intrinsic DNA curvature may depend strongly on the solution conditions (Sprous *et al.*, 1995; Dlakic & Harrington, 1996). For the nearest neighbour model, the roll and tilt angles of independent base-pair steps making up all 16 possible wedge angles have been estimated from experimental data and employed to determine sequence-dependent DNA bending energy for prediction of nucleosome positioning (Bolshoy *et al.*, 1991; De Santis *et al.*, 1992; Cacchione *et al.*, 1995). Therefore, the nearest neighbour models could provide important information for the understanding of DNA curvature.

4.1.3.2.2 DNA curvature: dynamic state studies

DNA is a long and flexible molecule that can be considered in constant motion under thermal perturbations. The static models (Travers, 1989; Steitz, 1990), discussed above, describe the DNA behaviour in time-average distribution and consider the ease of bending of particular sequences in purely static terms (Calladine & Drew, 1996).

Thermal motions effect the contour of the DNA helix axis, and produce substantial fluctuations in the angles between adjacent base-pairs. In the process of

nucleosome formation when the DNA is wrapped into the core particle, the DNA molecules must exhibit dynamic fluctuation. It is, therefore, reasonable to employ dynamic models to describe nucleosome positioning with respect to DNA curvature. Appreciation of this property has led to the development of models which attempt to describe the dynamic character of DNA curvature and account the stochastic contribution to the net curvature of the helix axis (Ulyanov & Zhurkin, 1984; Zhurkin, 1985; Griffith *et al.*, 1986; Srinivasan *et al.*, 1987; Hagerman & Ramadevi, 1990; Park & Breslauer, 1991; Olson *et al.*, 1993). For example, by comparison to the DNase I digestion properties of DNA sequences, an assay which is sensitive to the dynamic ability of sequence elements to adopt a bent conformation (Brukner *et al.*, 1995b), with nucleosome positioning information which is based on the static geometry of the individual elements (Satchwell *et al.*, 1986; Goodsell & Dickerson, 1994), Brukner *et al.* (1995a) have shown that a number of motifs such as TA elements and CCA/TGG elements, are more realistically described in the dynamic DNase I-based model. This may be due to the fact that the approach is appreciative of the ability of double-stranded DNA to bend towards the major groove in a dynamic situation. As discussed above, the nearest neighbour model more adequately describes the DNA curvature. Zhurkin *et al.* (1991) have argued that although some components remain undefined and unclear, the integration of the stochastic component into the static wedge angles model makes the stochastic wedge model more appropriate to describe DNA curvature and help to address nucleosome positioning with respect to the DNA bending propensity.

However, there is disagreement over this proposal. By determining the effect of dynamic fluctuations on physical manifestations of DNA curvature, De Santis *et al.* (1995) have argued that the main features of nucleosome positioning can be satisfactorily explained by a static curvature model. Clearly, further studies on these potentially important features of DNA structure and its role in nucleosome positioning need to be undertaken before a clear picture emerges.

4.1.4 Algorithms for nucleosome positioning

Although it has not yet fully understood exactly what the DNA sequence determinants of nucleosome formation are, there have been many methods developed for the prediction of preferred binding sites for the histone octamer on a defined DNA sequence.

4.1.4.1 Statistically based algorithms

Drew & Travers (1985) first noted that sequences such AAA or TTT prefer to be placed with their minor groove along the inside of a curve, whereas other sequences such as GGG, CCC and GGC prefer to have their minor groove along the outside of a curved DNA. By using statistical analysis of occurrences of di- or trinucleotides found in sequenced core particle DNA molecules, Satchwell *et al.* (1986) further identified that the AAA/TTT and AAT/ATT sequences have a marked preference to be positioned with their minor groove facing inwards towards the histone octamer, whereas GC tend to be found in the opposite phase and therefore facing away from the histone octamer. Their analysis also revealed the appearance of an out-of-phase 10 bp periodicity of A•T and G•C base-pairs in nucleosomal DNA (Drew & Travers, 1985; Satchwell *et al.*, 1986; Muyldermans & Travers, 1994). These sequence preferences and the periodicity of their occurrence constitute the essential determinants of rotational positioning.

The preference shown for the location of defined di- or trinucleotides in bent DNA correlates well with the preferred values of the roll angle for these dinucleotide steps (Calladine & Drew, 1986). Consequently, an algorithm for prediction of nucleosomal rotational positioning, mainly based on the correlation between a particular dinucleotide step and its location in nucleosome DNA, was developed and used for successful predictions (Calladine & Drew, 1986; Drew & Calladine, 1987). Since then, numerous statistical analyses of the profiles of periodic occurrences of di- or trinucleotides have been collected to provide evidence to support the occurrence of periodic signals in positioned nucleosomes (Pehrson, 1989; Lowman & Bina, 1990; Haran *et al.*, 1994; Bina, 1994; Muyldermans & Travers, 1994; Bolshoy,

1995). These include: (i) AA/TT occurring between base-pairs 1 and 56 from both ends of the core particle at an average spacing of 10.1 bp. This regular periodicity was not maintained between position 56 and the dyad axis (Satchwell *et al.*, 1986; Travers & Klug, 1987); (ii) Fourier analysis has shown that the periodic signals for AA/TT is 10.26 bp and for GG/CC is 10.0 bp, and that the AA or GG are dominant over TT or CC dinucleotides (Bina, 1994); (iii) repeated $(A/T)_3nn(C/G)_3$ motifs are particularly favourable for nucleosome formation (Shrader & Crothers, 1989). Although the rotational positioning of the DNA bound to the histone octamer is essentially the same in chromatosomes and nucleosome core particles, the periodic modulation of trinucleotide sequences in chromatosomes is not so regular or as pronounced as in core particle DNA (Muyldermans & Travers, 1994). Moreover, the signal NGGR (where N=A, T, G, C; R=A or G), which is frequently found to be located asymmetrically within chromatosomal DNA, and is unique to chromatosome DNA, may play a direct role in positioning chromatosomes in condensed chromatin (Travers & Muyldermans, 1996).

Information accumulated from the study of periodic signals in DNA has been employed to produce algorithms for predicting nucleosome positioning. However, in eukaryotic DNA the nucleosome positioning pattern appears relatively weak, redundant and still not fully understood. Recently, alignment procedures have been developed for the extraction of positioning patterns: (i) multiple alignment algorithms based on a statistical matching of the sequences, can be aligned to reveal the hidden patterns in DNA sequences (Ioshikhes *et al.*, 1992; Ioshikhes *et al.*, 1996; Bolshoy *et al.*, 1996); (ii) multi-alphabet consensus algorithm based on a binomial statistic can be used to evaluate a pattern's occurrence among a set of nucleotide sequences (Ulyanov & Stormo, 1995); and (iii) hidden Markov models are based on an ability to describe a series of observations by a hidden stochastic process (Borodovsky & Peresetsky, 1994; Baldi *et al.*, 1996). Although these alignment procedures may help to identify and find the periodic signals in the DNA, they are, as yet, far from reliable methods for predicting nucleosome positioning in general.

4.1.4.2 Structurally based algorithms

Many studies have demonstrated a relationship between nucleosome positioning and sequence-directed curvature ((Drew & Traver, 1985; Stachwell *et al.*, 1986; Neubauer *et al.*, 1986; Drew & Calladine, 1987; Travers & Klug, 1987; Kefalas *et al.*, 1988; Travers, 1989; Pennings *et al.*, 1989; Costanzo *et al.*, 1990). Recently, structurally based algorithms have been developed for prediction of nucleosome translational positioning. All of these are fundamentally based on the nearest neighbour wedge models which mainly takes the roll and tilt angles between base steps into account. Their capacity to successfully predict positioning mainly attributes to the determination of the 16 wedge angles which contribute to the curvature (Bolshoy *et al.*, 1991), and the results suggest that the intrinsic bendability of DNA is very important for nucleosome positioning (Shpigelman *et al.*, 1993; De Santis *et al.*, 1993; Sivolob & Khrapunov, 1995; Cacchione *et al.*, 1995; De Santis *et al.*, 1995; Kropp *et al.*, 1995; Kralovics *et al.*, 1995). However, as argued before, nucleosome formation is a dynamic process rather than a static situation. Therefore, incorporation of dynamic components into the algorithms may be more useful and will provide more precisely predicted nucleosome positioning information.

4.1.5 Contribution of linker histones to nucleosome positioning

Linker histones can help the folding of chromatin into higher order structure and studies have indicated that this kind of higher order structure can modulate nucleosome positioning (Thoma *et al.*, 1979; Thoma & Koller, 1981). By implication, the linker histones are thought to be capable of determining nucleosome positioning. Indeed, evidence has suggested that removal of linker histone from chromatin can result in the migration of the histone octamer (Satchwell *et al.*, 1986; Satchwell & Travers, 1989). However, the actual mechanism is not well defined.

Linker histone H1 and its variant H5 have been regarded as repressive components in gene expression by virtue of the fact that they associate with positioned nucleosomes and help them to fold into higher order chromatin structure.

The association between linker histones and positioned nucleosomes is very dynamic (Caron & Thomas, 1981; Thomas & Rees, 1983). The dynamic nature of this interaction has relevance in gene activation as the weak interaction between linker histones and nucleosomes allows the linker histones or the core particle octamer to be replaced or removed (Hill *et al.*, 1991; Cote *et al.*, 1994; Wilson *et al.*, 1996) allowing the sliding and transient dissociation of nucleosomes along the DNA fragment.

4.1.6 Aim of this chapter

Chromatin structure can play a decisive role in the mechanism of gene regulation. Many studies have shown that nucleosomes can prevent transcription by virtue of their locations within the gene regulatory region (Hill *et al.*, 1991; Cote *et al.*, 1994; Wilson *et al.*, 1996). For example, the chicken adult β-globin gene remains transcriptionally silent early in development even when its 3′ enhancer is activating embryonic epsilon-globin gene expression (Hesse *et al.*, 1986; Choi & Engel, 1986). It has been suggested that local chromatin structure may play a role in this process (Buckle *et al.*, 1991). Furthermore, the binding of linker histone to the nucleosome helps to form a more compact higher order chromatin structure, providing less accessibility for transcription factors and influencing the overall control of gene expression. Many studies have shown that linker histone globular domain has major potential to influence the binding or the location of binding of core histones to the DNA by virtue of its role as an agent of higher order folding (Simpson, 1978; Allan *et al.*, 1980; Thoma *et al.*, 1983; Allan *et al.*, 1986). However, whether the binding of a linker histone molecule to an individual nucleosome can alter its position is unclear.

As described above, previous studies have shown that the characteristics of the DNA sequence make a major contribution to positioning nucleosome. To understand the contribution of sequence-directed nucleosome positioning to the mechanism of gene expression, it is important to ascertain nucleosome positioning

with respect to a whole gene sequence. Most studies of sequence-directed nucleosome positioning have analysed only short regions removed from their natural sequence context and have focused on regulatory DNAs. Recently, work in this laboratory has developed a technique – monomer extension - which enables sequence-directed nucleosome translational positioning to be mapped over extensive stretches of DNA, at high resolution and quantitatively (Yenidunya *et al.*, 1994). This technique has already been applied to produce a long-range, high resolution positioning map for the entire chicken adult β-globin gene region (Davey *et al.*,1995), and to assess the effects of DNA methylation on positioning within this promoter (Davey *et al.*, 1997).

In this study, monomer extension has been used to map chromatosome-positioning sites in a 1.5 kb region of the chicken adult β-globin gene. The capacity of monomer extension to map nucleosomes at high resolution and long DNA sequence enables the influence of linker histones to be investigated. Here, therefore, the aim was to understand how the linker histone modulates nucleosome positioning, to investigate their affinity for positioned nucleosomes and to elucidate any differences in these contexts between linker histone subtypes. The results have provided insights into all these issues and have suggested the binding mode of linker histone globular domain to positioned nucleosomes.

4.2 Materials and methods

4.2.1 Materials

M9 salt – 5 X M9 salt contained 0.24 M Na_2HPO_4, 0.11 M KH_2PO_4, 0.043 M NaCl, and 0.093 M NH_4Cl.

M9 minimal medium – 1000 ml of M9 minimal medium was made by dissolving 15 g of agar in 770 ml water and autoclaving it. After autoclaving, cool to 55 °C, then add 200 ml of 5 X M9 salt, 2 ml of 1 M MgSO4, 20 ml of 20 % glucose, 100 µl of 1M $CaCl_2$, 1ml of 1 mg/ml thiamine and 1000 µl of 100 mg/ml ampicillin, and then pour it into plate.

SOB – 500 ml SOB was prepared by dissolving 10 g tryptone, 2.5 g yeast extract, 0.25 g NaCl and 5 ml of 250 mM KCl in 470 ml water, adjust the pH 7.0, adjust volume to 492.5 ml, and then autoclave it. Add 2.5 ml of 2 M $MgCl_2$ and 5 ml of 1 M $MgSO_4$ just before use.

TBG – Dissolve bacto tryptone at 1.2%, yeast extract at 2.4%, glycerol at 0.4%, KH_2PO_4 at 17 mM, and K_2HPO_4 at 55 mM in water and autoclave it. Add glucose to 20 mM before use.

4.2.2 Phagemid construction

Appropriate DNA fragments of the chicken β-globin gene promoter region were cloned from plasmid pCARB 4.4 (Walmsley *et al.*, 1991) into the polylinker of pBluescript phagemid vectors (Stratagene). All eight overlapping fragments were cloned into pBluescript KS(-) in both orientations by using one or two blunt cutting sites to generate the following phagemids for monomer extension: EcXmV (-1052 to +426; relative to the cap site of the β-globin gene), XmEcS (-1052 to +426), SmEc (-

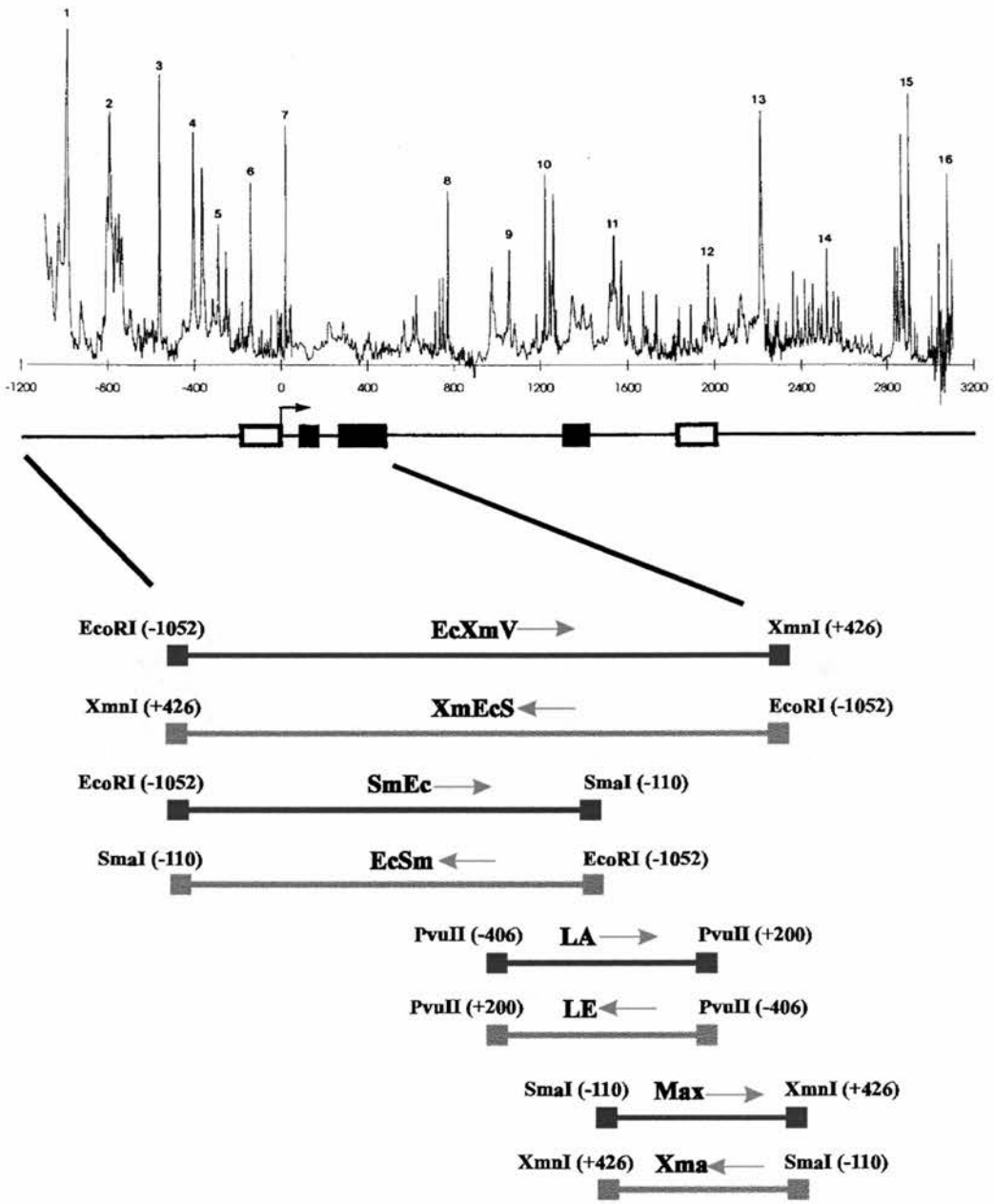


Figure 4-2 Map of the histone octamer positioning sites for the chicken adult β-globin gene region (taken from Davey *et al.*, 1995). Below the map are shown the location and orientation of the globin sequences contained within the mapping constructs employed in this study. The arrows on the mapping constructs indicate the 5'→3' orientation of the globin sequences.

1052 to –110), EcSm (-1052 to –110), LA (-406 to +200), LE (-406 to +200), Max (-110 to +426), and Xma (-110 to +426) (Figure 4-2). All phagemids were transformed into *E.coli* strain DH11S for preparation of phagemid single-stranded DNA.

4.2.3 Preparation of M13KO7 helper phage stock

5 ml TYP (1.6 % bacto tryptone, 1.6 % yeast extract, 0.5 % NaCl and 0.25 % K_2HPO_4) were inoculated with 100 µl DH11S M9 glycerol and incubated at 37 °C for 6 hours (mid to late log phase). 200 µl of log phase DH11S were mixed with 10 to 20 µl of M13KO7 phage (Pharmacia) 10-fold serial dilutions from 10^{11} pfu/ml to 10^3 pfu/ml and incubated at room temperature for 5 minutes. 50 or 100 µl aliquots of each mix were spreaded on dried TYP plates (1.6 % bacto tryptone, 1.6 % yeast extract, 0.5 % NaCl, 0.25 % K_2HPO_4 and 1.5 % agar) containing 70 µg/ml kanamycin and incubated at 37 °C over night. Individual, well-isolated colonies were picked and mixed with 35 ml of TYP containing 70 µg/ml kanamycin at 37 °C for 8 to 12 hours. The culture was spun for 20 minutes at 3.5 K to pellet most of the bacteria. The supernatant was respin at 10 K for 20 minutes. This second supernatant was incubated at 55 °C for 45 minutes to kill residual bacteria and then stored at 4 °C.

4.2.4 Preparation of DH11S phagemid SOB/glycerol stock

All phagemids were transformed into *E.coli* DH11S strain and LB/glycerol stocks made as described in Chapter 2. The phagemid LB/glycerol stock was spread on LB agar (amp) and incubated at 37 °C over night. Colonies were picked and spread on M9 plates (amp). After 48 hours of incubation at 37 °C, colonies were inoculated into SOB for over night growth. An aliquot of over night culture was mixed with an equal volume of SOB/glycerol (60%:40%) and stored at –70 °C.

4.2.5 Preparation of single-stranded DNA

E.coli DH11S (F+) cells carrying the recombinant pBluescript-globin plasmids were infected with M13KO7 helper phage to prepare the single-stranded DNA. 100 ml TBG was inoculated with 300 µl DH11S SOB plasmid glycerol and 500 µl of M13KO7 helper phage (~ 10^{11} pfu/ml), and incubated at 37 °C. After 90 minutes incubation, 100 µg/ml of ampicillin and 70 µg/ml of kanamycine were added, and the culture grown for a further 12 to 14 hours.

Phage particles were isolated by spinning the culture twice at 8K for 20 minutes at 4°C. The 2° supernatant was added to 0.25 volume of 40 % PEG6000 in 2.5 M NaCl and incubated for 60 minutes on ice to precipitate the phage particles. The phage particle was collected by spinning at 8K for 30 minutes at 4 °C. The pellet was further spun at 3K for 5 minutes at 4°C and the residual PEG/NaCl was removed completely. The pellet was resuspend in 400 µl TE and split into two 200 µl aliquots. 1 ml of Tri Reagent (Sigma) was added to each aliquot and the mixture incubated at room temperature for 15 minutes with occasional vortex, 200 µl of chloroform was added and the mixture was incubated at room temperature for a further15 minutes with occasional vortex. After this incubation, the mixture was spun at 13 K for 15 minutes and the colourless aqueous phase removed (containing RNA). The remaining interphase and organic phase contained single-stranded DNA and protein. To precipitate the DNA, 0.3 ml of ethanol was added and the mixture incubated at room temperature for 15 minutes. The DNA pellet was collected by spinning at 6.5K for 5 minutes, all residual supernatant removed, and the pellet washed twice with 1 ml of 0.1 M sodium citrate/10 % ethanol, then twice with 70 % ethanol. The dried pellet was resuspended in 675 µl TE and then EDTA added to 2.5 mM, SDS added to 0.5%, and NaAc pH 5.2 added to 0.3 M followed by extraction with three to four times phenol and once with chloroform. After ethanol precipitation, the DNA was resuspend in 500 µl TE and incubated on ice for 60 minutes with 300 µl 20 % PEG6000 in 2.5 M NaCl. After this, pellet was collected by spinning the mixture at 13K for 15 minutes followed by two washes, with 70 % ethanol and resuspension in 210 µl TE. The quality of single-stranded DNA was

assessed by running two 5 µl samples, one incubated at room temperature and the other at 95 °C for 4 minutes, on a 1.6 % agarose gel.

4.2.6 Standard *in vitro* nucleosome reconstitution

Core particle DNA was prepared from pCBA4.4 which comprises a 4.4-kb *Eco*RI-*Bam*HI fragment of chicken β-globin sequence (-1052 to +3369, relative to the transcription start site of the adult gene) in pBluescript KS (Davey *et al.*, 1995). Reconstitution was carried out at a core histone/DNA ratio of 0.5:1 (w/w). The reconstitution mixture contained 25 µg of pCBA4.4 DNA (linearised by digesting with BsaHI), 12.5 µg of chicken erythrocyte core histones, 2 M NaCl, 1 X TE, and 0.1 mM PMSF. After incubation at room temperature for 5 to 10 minutes, the reconstitution was dialysed against a linear 2 M to 0.4 M NaCl gradient (in 1 X TE with 0.1 mM PMSF) at 4°C for 3 to 4 hours. After this, the reconstitutes were further dialysed against 80 mM NaCl (in 1 X TE with 0.1 mM PMSF) at 4°C overnight. The samples were then collected and stored at 4°C until use.

4.2.7 Titration of GH1/GH5 to reconstituted chromatin, Micrococcal nuclease digestion and preparation of monomer DNA

GH1 or GH5 was added to the reconstituted chromatin at 1:1 molar ratio and incubated on ice for 60 minutes. After this, micrococcal nuclease (MNase) was added to the mixture to 5 unit/ml and $CaCl_2$ to 1 mM and then the digestion was proceeded by incubating the mixture on ice for 30 minutes followed by 105 seconds at 37 °C for trimming of the digestion products. Finally, EDTA was added to 15 mM to stop the digestion reaction. In the MNase digestion time course experiment, the preparation methods were the same as above, only the trimming durations were 0, 35, 70, 105, 140 and 300 seconds.

DNA purified from micrococcal nuclease digestion was electrophoresed in either a native 7 % polyacrylamide gel or 4.5 % metaphor gel (FMC) to achieve a significant separation between core particle DNA (~ 146 bp) and chromatosomal DNA (~ 168 bp). The correctly sized DNAs were excised from the gel and removed by electroelution from a dialysis bag. Electroeluted DNA was purified by phenol:chlorofom and chloroform:isoamyl alcohol extractions and sodium acetate/ethanol precipitation. The DNA pellet was washed with 70 % ethanol and then resuspend in an appropriate volume of TE. Typically, a 25 µg plasmid reconstitute yielded 1.5 – 2.0 µg of monomer DNA in total (both core particle DNA and chromatosomal DNA).

4.2.8 *In vitro* nucleosome reconstitution at 37 °C

Reconstitution was carried out at a core histone/DNA ratio of 0.5:1 (w/w). The reconstitution mixture, which again contained linear pCBA4.4 DNA, chicken erythrocyte core histones, 2 M NaCl, 1 X TE and 0.1 mM PMSF, was incubated at room temperature for 5 to 10 minutes and then dialysed against a linear 2 M to 0.4 M NaCl gradient in 1 X TE at 4°C for 3 to 4 hours. After this, the reconstitute apparatus was moved to a warm room where the temperature was 35-37 °C, and GH1 or GH5 added at the ratio of 1:1 to the core histone concentration. The reconstitution was carried out by dialysis against a linear 0.5 M to 80 mM NaCl gradient (in 1 X TE with 0.1 mM PMSF) for 3 to 4 hours at 37 °C. Finally, the reconstitution was further dialysed against 80 mM NaCl (in 1 X TE with 0.1 mM PMSF) for another 1 hour at 37 °C. The samples were collected and stored at 4 °C until use.

4.2.9 Labelling and alkali-denaturation of the monomer DNA

The labelling method used was based on the 5′ protruding end forward reaction described by Maniatis *et al.* (1982). In a 15 µl reaction, monomer DNA was mixed with 1 X 5′-end labelling buffer (50 mM Tris pH 7.6, 10 mM $MgCl_2$, 5 mM

DTT, 0.1 mM spermidine and 0.1 mM EDTA), 3 µl of [γ-^{32}P]ATP (3000 Ci/mmol) and 10 units of T4 polynucleotide kinase (Amersham), and the reaction mix incubated at 37 °C for 60 minutes followed by phenol:chlorofom and chloroform:isoamyl alcohol extractions. The labelled DNA was separated from unincorporated label by spinning through a 1 ml Sephadex G50 column equilibrated with water. The total volume of eluent was adjusted to 54 µl. 6 µl of 2 M NaOH was added and the mix incubated at room temperature for 10 minutes (Hattori & Sakaki, 1986). The labelled DNA was neutralised with 24 µl 5M NH_4Ac pH 7.5 and precipitated with ethanol. The DNA pellet was washed with 70 % ethanol and then resuspended in TE at 10 ng/µl.

4.2.10 Monomer extension reactions

A schematic outline of the monomer extension procedure has been shown in Figure 4-3. A 20 to 50 ng sample of freshly denatured monomer DNA was mixed with 0.8 to 1.5 µg of the appropriate single-stranded DNA (according to the size of the single-stranded DNA) in 100 mM NaCl, 2 mM DTT, 20 mM Tris pH 7.5 and 20 mM $MgCl_2$ to a total volume of 25 µl; annealing was effected by denaturation at 95 °C for 3 minutes, the temperature dropped to 80 °C for 30 seconds, and then gradually lowered to 55 °C over 45 minutes. After annealing, the reaction mix was diluted to 50 µl by the addition of dNTPs (10 µM final concentration of each dATP, dCTP, dTTP and dGTP), BSA (100 µg/ml final concentration), 5 units of *E. coli* DNA polymerase (Klenow fragment; Amersham), 20 units of appropriate restriction enzyme and water, and incubated at 37 °C for 60 minutes followed by phenol:chlorofom and chloroform:isoamyl alcohol extractions and sodium acetate/ethanol precipitation. The product was washed with 70 % ethanol and then resuspend in 8 to 10 µl of sequence stop mix. Samples were heat denatured and then electrophoresed in 6 % denaturing polyacrylamide gels with size standards of long-geared C and T ^{35}S sequencing reactions of M13mp18 ssDNA (USB) and ^{32}P end-labelled HinFI and DdeI digests of phage lambda.

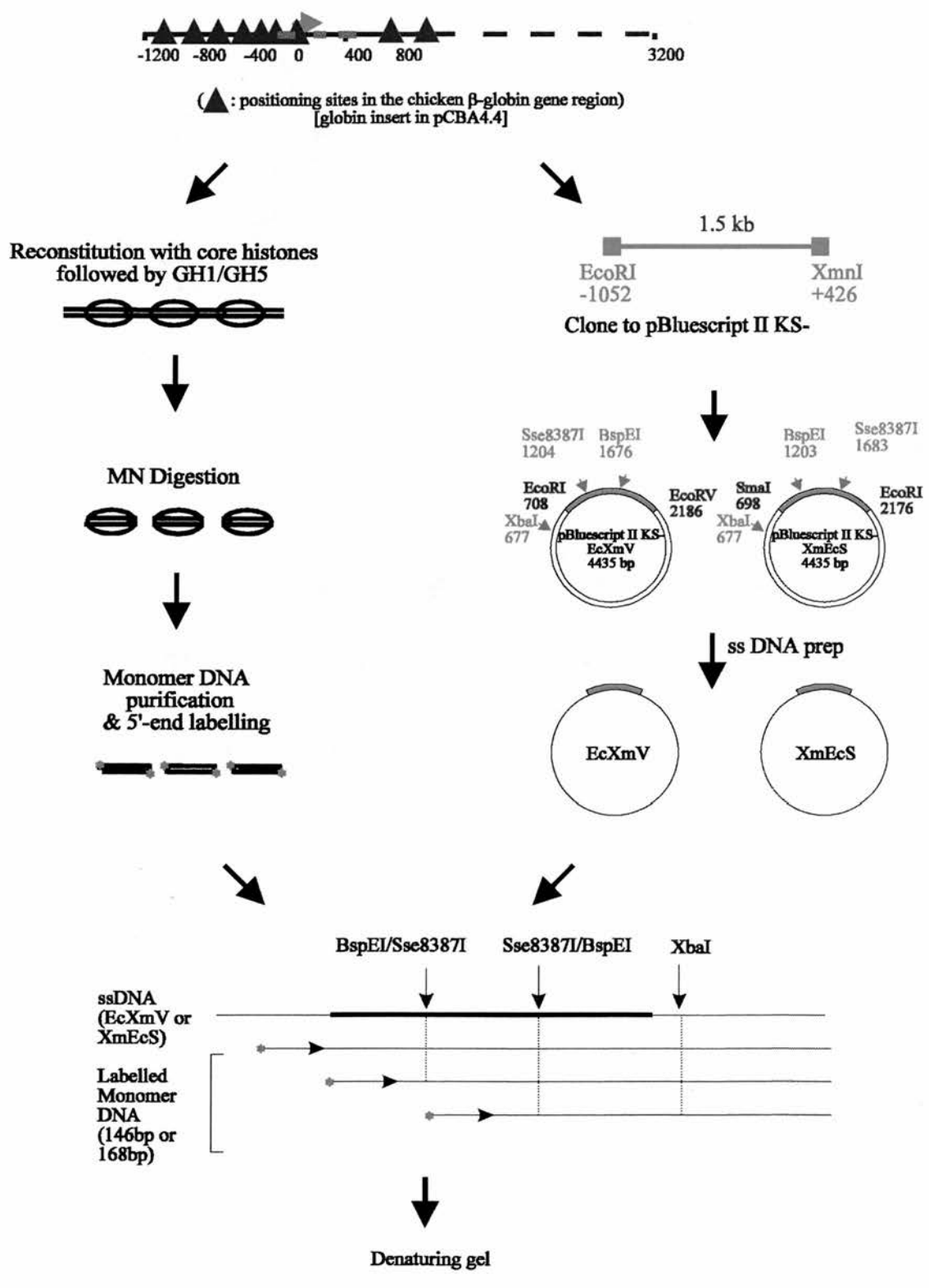


Figure 4-3 Schematic outline of the monomer extension procedure.

4.2.11 Analysis of the nucleosome mapping from 6 % denaturing polyacrylamide gels

Quantitative densitometer scans were obtained for each extension reaction after PhosphorImager (Molecular Dynamics) analysis of the dried gels. The equation which converted from coordinates of markers to DNA size was determined by a six or higher order polynomial (Figure 4-4). The correlation coefficient is greater than 99.9 %. The lengths of extension products were converted by the application of the equation after densitometry of phosphor images. The locations of the positioning site boundaries were determined by the lengths of extension products with respect to a unique restriction enzyme cutting site in the β-globin gene. Densitometry traces carried out in the absence of a restriction enzyme were considered as background for quantitation adjustment. Normalisation was determined by common nucleosome positioning sites within vector sequence.

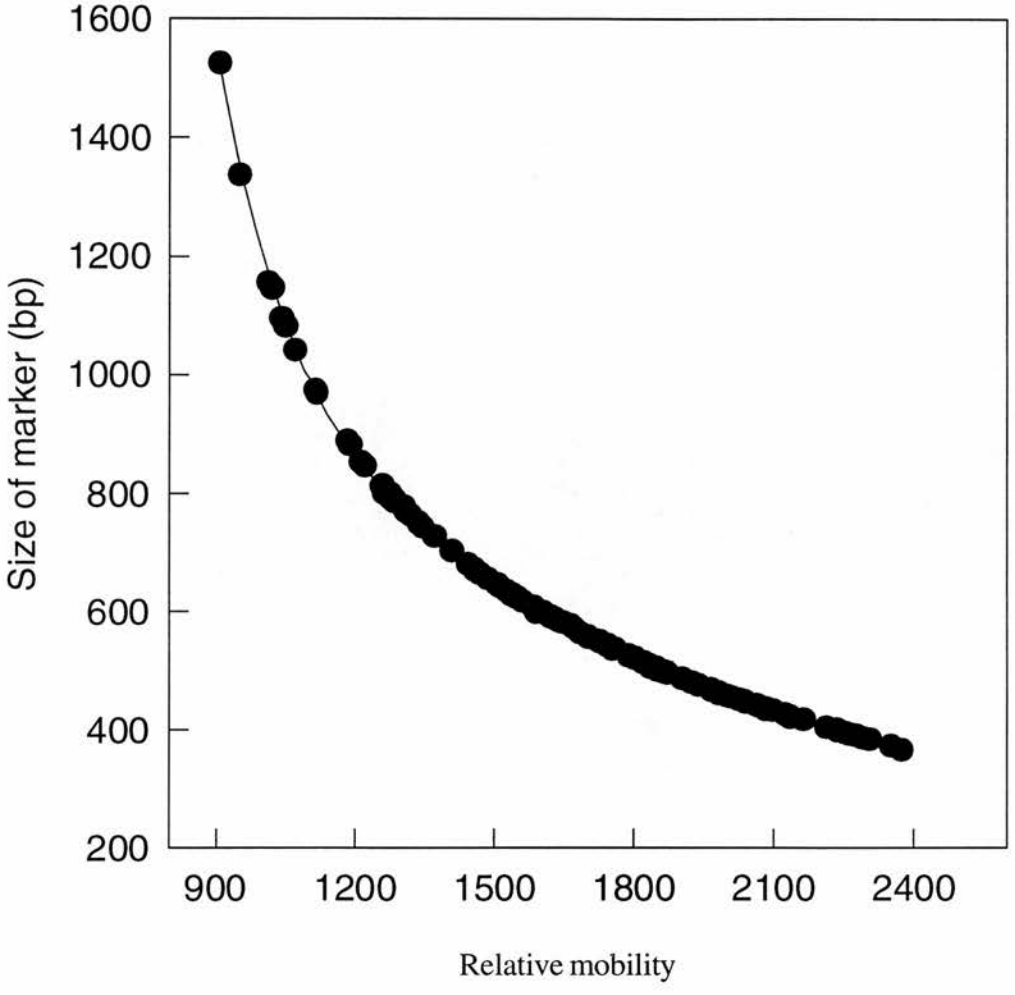


Figure 4-4 Correlation between DNA mobility and DNA size derived from analysis of restriction and sequencing markers on a 6% denaturing polyacrylamide gel.

4.3 Results

Nucleosomes often adopt precise positions with respect to their underlying DNA sequence (Simpson, 1991; Thoma, 1992). Previously, Davey *et al.* (1995) have applied the technique of monomer extension to produce a high resolution, long-range (4.4 kbp), translational positioning map for the chicken adult β-globin gene *in vitro*. In this study, the approach was essentially similar to monomer extension except that the addition of linker histones or linker histone globular domains was incorporated into the reconstitution procedure. This enabled the isolation of a population of chromatosomal DNA (168 bp) in addition to the core particle DNA for subsequent mapping. Using this approach, it was possible to identify and compare the effects of linker histones or linker histone globular domains on nucleosome positioning.

4.3.1 Preparation of globular domains of linker histones H1 (GH1) and H5 (GH5)

4.3.1.1 Expression and purification of recombinant GH1 and GH5

pGH1-41, the GH1 expression construct employed (from Chapter 3), contained an inframe H1 fragment encoding a 76 amino acid GH1 peptide with an additional N-terminal methionine residue (Figure 4-5). pGH5, the GH5 expression construct employed, contained an inframe H5 fragment encoding a 79 amino acid GH5 peptide with an extra N-terminal methionine residue (Figure 4-6). In both cases the expression vector was the T7 polymerase based plasmid pT79 (see Chapter 3.2.1 for details). *E.coli* (DE3) containing pGH1-41 or pGH5 was grown in LB and induced in mid-log phase by the addition of IPTG. After 3 hours induction, the bacteria were collected, lysed, and the bacterial debris was removed by centrifugation. The soluble phase was then extracted with 5% perchloric acid, conditions under which the GH1 and GH5 are exclusively soluble (see Chapter 2.14 for details). The recombinant GH1 and GH5 peptides were then recovered from the

```
 1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
ATG CCA GCA GGT CCA AGC GTC ACC GAG CTG ATC ACC AAG GCC GTG
Met Pro Ala Gly Pro Ser Val Thr Glu Leu Ile Thr Lys Ala Val

16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
TCC GCC TCC AAG GAG CGC AAG GGG CTC TCC CTC GCC GCG CTC AAG
Ser Ala Ser Lys Glu Arg Lys Gly Leu Ser Leu Ala Ala Leu Lys

31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
AAG GCG CTT GCC GCC CGC GGC TAC GAC GTG GAG AAG AAC AAC AGC
Lys Ala Leu Ala Ala Gly Gly Tyr Asp Val Glu Lys Asn Asn Ser

46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
CGC ATC AAG CTG GGG CTC AAG AGC CTC GTC AGC AAG GGC ACC CTG
Arg Ile Lys Leu Gly Leu Lys Ser Leu Val Ser Lys Gly Thr Leu

61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
GTG CAG ACC AAG GGC ACC GGC GCC TCG GGC TCT TTC AAG CTG AAT
Val Gln Thr Lys Gly Thr Gly Ala Ser Gly Ser Phe Lys Leu Asn

76  77
AAA AAG
Lys Lys
```

Figure 4-5 The coding sequence of recombinant GH1.

```
 1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
ATG TCG GCA TCG CAC CCC ACC TAC TCG GAG ATG ATC GCG GCG GCC
Met Ser Ala Ser His Pro Thr Tyr Ser Glu Met Ile Ala Ala Ala

16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
ATC CGT GCG GAA AAG AGC CGC GGC GGC TCC TCG CGG CAG TCC ATC
Ile Arg Ala Glu Lys Ser Arg Gly Gly Ser Ser Arg Gln Ser Ile

31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
CAG AAG TAC ATC AAG AGC CAC TAC AAG GTG GGC CAC AAC GCC GAT
Gln Lys Tyr Ile Lys Ser His Tyr Lys Val Gly His Asn Ala Asp

46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
CTG CAG ATC AAG CTC TCC ATC CGA CGT CTC CTG GCT GCC GGC GTC
Leu Gln Ile Lys Leu Ser Ile Arg Arg Leu Leu Ala Ala Gly Val

61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
CTC AAG CAG ACC AAA GGG GTC GGG GCC TCC GGC TCC TTC CGC TTG
Leu Lys Gln Thr Lys Gly Val Gly Ala Ser Gly Ser Phe Arg Leu

76  77  78  79  80  81
GCC AAG AGC GAC AAG TGG
Ala Lys Ser Asp Lys Trp
```

Figure 4-6 The coding sequence of recombinant GH5.

acid-soluble phase by acetone precipitation. As shown in Figure 4-7, the recombinant proteins were of high purity. Purified GH1 and GH5 peptides were produced at reasonably high yields (44 and 16 mg/L, respectively).

4.3.1.2 Chromatin reconstitution and chromatosome protection with recombinant GH1 and GH5

Chromatosome protection is a unique functional capacity of the linker histone globular domain (Allan *et al.*, 1980). Recombinant GH1 and GH5 were therefore tested for their chromatosome protection ability. *In vitro* reconstitution of core histones onto DNA (pCBA4.4) was carried out by a standard procedure (see Chapter 4.2.6). The resulting reconstituted chromatin was then titrated with recombinant GH1 or GH5 and digested with MNase. The results of a typical reconstitution/chromatosome protection experiment are presented in Figure 4-8. Digestion of chromatin reconstituted without GH5 resulted in protection of only 146 bp DNA fragments (Figure 4-8, lane 1). Reconstitution with increasing molar ratios of recombinant GH5 revealed a corresponding increase in the degree to which 168 bp DNA fragments were protected (Figure 4-8, lane 2 to 9). This indicates that recombinant GH5 was functionally equivalent to its native counterpart in respect of this fundamental structural assay. Similar results were obtained with GH1 (data not shown).

In order to optimise digestion conditions for the production of chromatosomal length DNA (168 bp) (aiming for equivalent amounts of 168 and 146 bp DNA), the effects of MNase trimming during enzyme digestion were investigated. Reconstituted chromatin was titrated with 1 molecule of recombinant GH1 or GH5 per nucleosome and then digested with MNase. Digestion was for the standard 30 minutes on ice, followed by increasing periods of incubation at 37 °C when trimming takes place. As shown in Figure 4-9, trimming time strongly influences chromatosomal DNA protection for both the GH1 and GH5 reconstitutes. Comparison of the GH1 and GH5 digests show that both proteins gradually lost their protective ability as the trimming time increased. On the basis of these experiments,

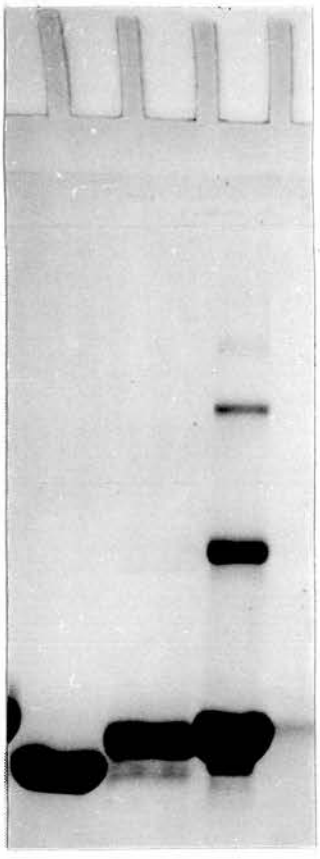


Figure 4-7 15% SDS gel analysis of purified, recombinant GH1 (lane 1) and GH5 (lane 2). Trypsin prepared wild-type GH5 is shown in lane 3).

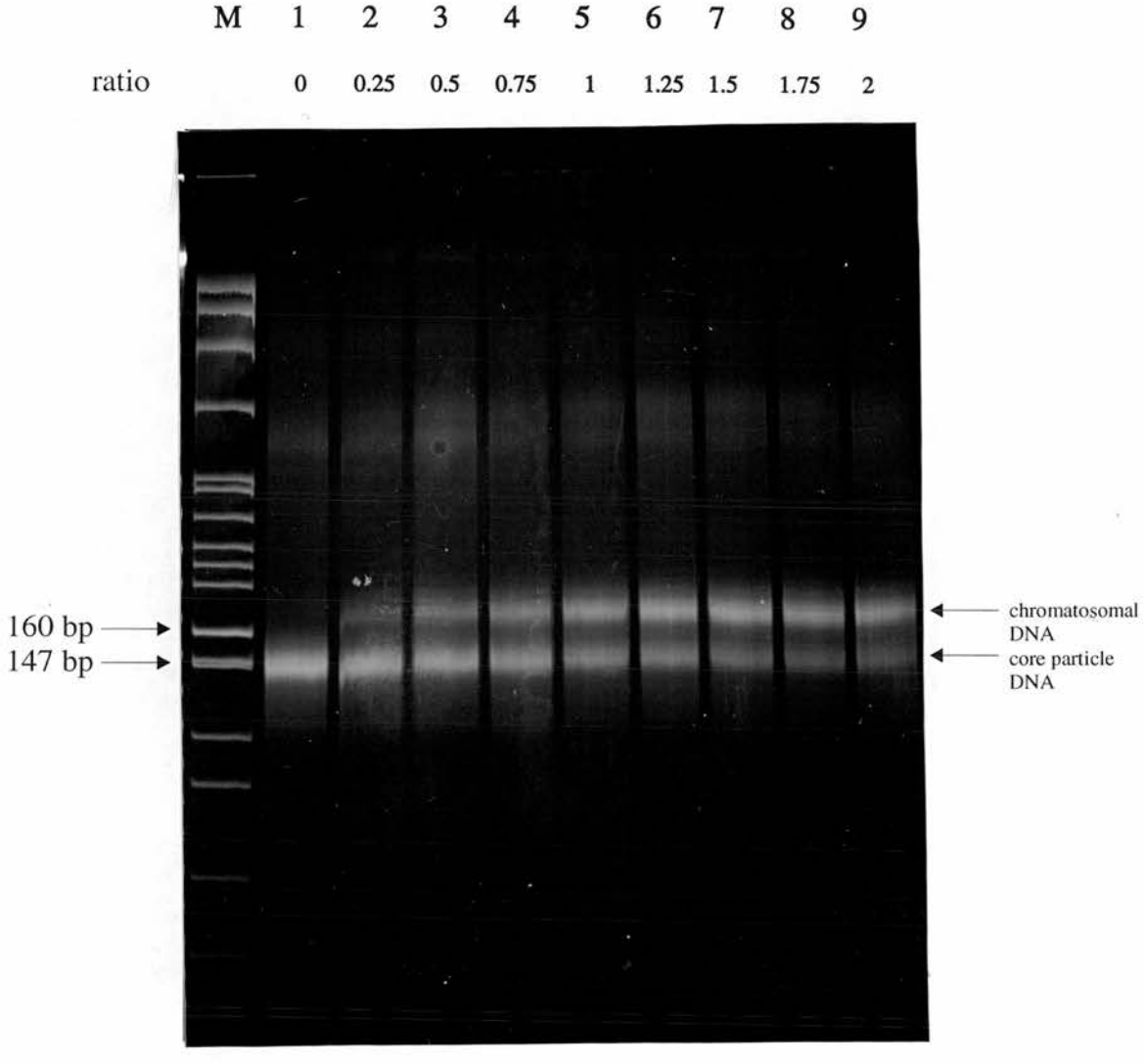


Figure 4-8 Chromatosome protection of reconstituted chromatin containing increasing amounts of GH5. DNAs purified from MNase digested chromatin were run in a 6% polyacrylamide gel. The GH5 to core histone octamer ratio is indicated above each lane. The marker (M) was an MspI digest of pBR322 DNA.

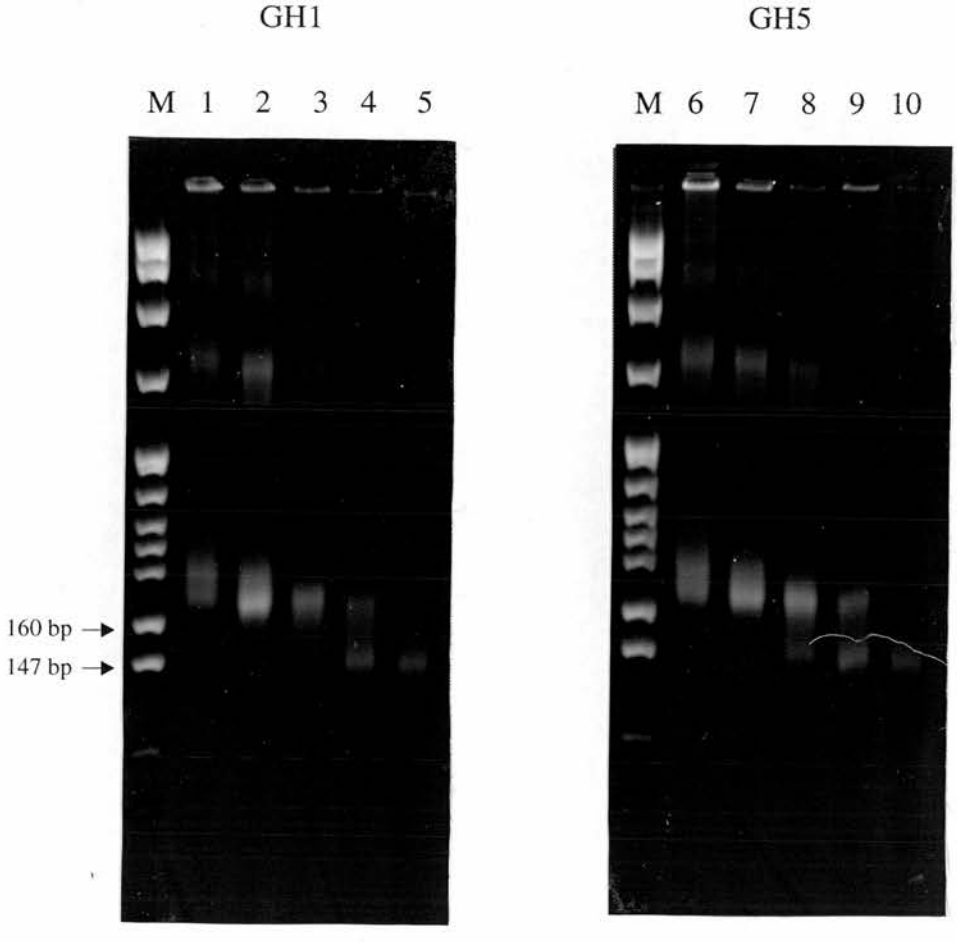


Figure 4-9 Chromatosome protection of reconstituted chromatin containing GH1 or GH5 as a function of MNase trimming time. DNAs purified from MNase digested chromatin were run in a 4.5% agarose gel. GH1 or GH5 reconstitutes are trimmed for 0 second (lane 1 & 6), 35 seconds (lanes 2 & 7), 70 seconds (lane 3 & 8), 105 seconds (lanes 4 & 9) and 140 seconds (lanes 5 & 10), respectively. The marker (M) was an MspI digest of pBR322 DNA.

it was decided to employ a MNase trimming period of 105 seconds, a condition which provided almost equivalent amounts of 146 and 168 bp DNA.

4.3.2 Monomer extension: determining core particle and chromatosome boundaries

This laboratory has developed a method (monomer extension) to map the precise translational positions adopted by core histone octamers reconstituted onto long DNA sequences (Yenidunya *et al.*, 1994). The nucleosome positioning sites for the entire chicken adult β-globin gene have been mapped by this method (Figure 4-2) (Davey *et al.*, 1995). This analysis demonstrated that strong positioning sites display a notable periodicity of ≈ 200 bp, particularly at the 5′ end of the gene (sites 1-7, Figure 3 in Davey *et al.*, 1995, Figure 4-2). As these sites flank the coding sequence they could be involved in packaging the DNA into higher-order chromatin.

Although DNA sequence contributes to determining nucleosome positioning, other factors may modulate this behaviour. One of the most abundant chromosomal proteins in the nucleus, apart from the core histones, are the linker histones. To what extent these proteins influence nucleosome positioning has been investigated in only a few special instances (Pennings *et al.*, 1991; Hayes *et al.*, 1993; Pennings *et al.*, 1994; Ura *et al.*, 1995). However, there has been no systematic study on how linker histones effect the positioning of a large population of positioning sites. Also, in this context, no detailed comparison has been made of H1 with H5 or of GH1 with GH5. In an attempt to study these issues, the unique comparative properties of the monomer extension technique have been exploited.

4.3.2.1 Monomer extension – Preparation and characterisation of core particle and chromatosomal DNA populations

A schematic outline of the monomer extension procedure used to map chromatosome binding sites adopted *in vitro* is shown in Figure 4-3. A plasmid containing the chicken adult β-globin gene, its enhancer, and flanking sequences

(pCBA4.4; Figure 4-2) was reconstituted with limiting amounts of core histones. At the ratio employed there was approximately 1 core histone octamer per 500 bp of DNA. Aliquots of this reconstituted chromatin were then titrated with one molecule of GH1/GH5 per core histone octamer. As a control, some of the chromatin was treated exactly the same as the GH1/GH5 reconstitute but no GH1/GH5 was added. These chromatins were digested with MNase under the conditions established above, to produce populations of chromatosome and core particle DNAs. Chromatosomal and core particle DNAs were fractionated in either a 7% polyacrylamide gel or a 4.5% agarose gel. Bands, identified after ethidium bromide staining, were cut from the gel and recovered by electroelution. Purified chromatosomal and core particle DNA molecules were 5′ end-labelled and used as heterogeneous populations of primers for extension on each of a set of 8 single-stranded phagemids containing overlapping sections of the 1.5-kb globin region (Figure 4-2).

To assess the quality of core particle and chromatosomal DNA, they were run on a 4.5% agarose gel. As shown in Figure 4-10A, core particle and chromatosomal DNAs comprised relatively pure populations of 146 and 168 bp DNAs, respectively. There was no substantial difference between core particle DNAs isolated from a core histone:DNA reconstitute and reconstitutes which also contained GH1 or GH5. Furthermore, there was no difference between chromatosomal DNA isolated from GH1- or GH5-containing chromatin.

5′ end-labelled DNAs were also analysed in a 6% denaturing polyacrylamide gel. In this analysis, DNAs are studied in a single-stranded state and this tests whether core particle or chromatosomal DNA has sustained significant levels of internal nicking during preparation. The result, presented in Figure 4-10B, shows that both chromatosomal and core particle DNA are predominately full length 168 and 146 bp, respectively (this autoradiograph was deliberately overexposed to attempt to identify nicked DNA). Core particle and chromatosomal DNAs from all subsequent experiments were routinely analysed by this approach.

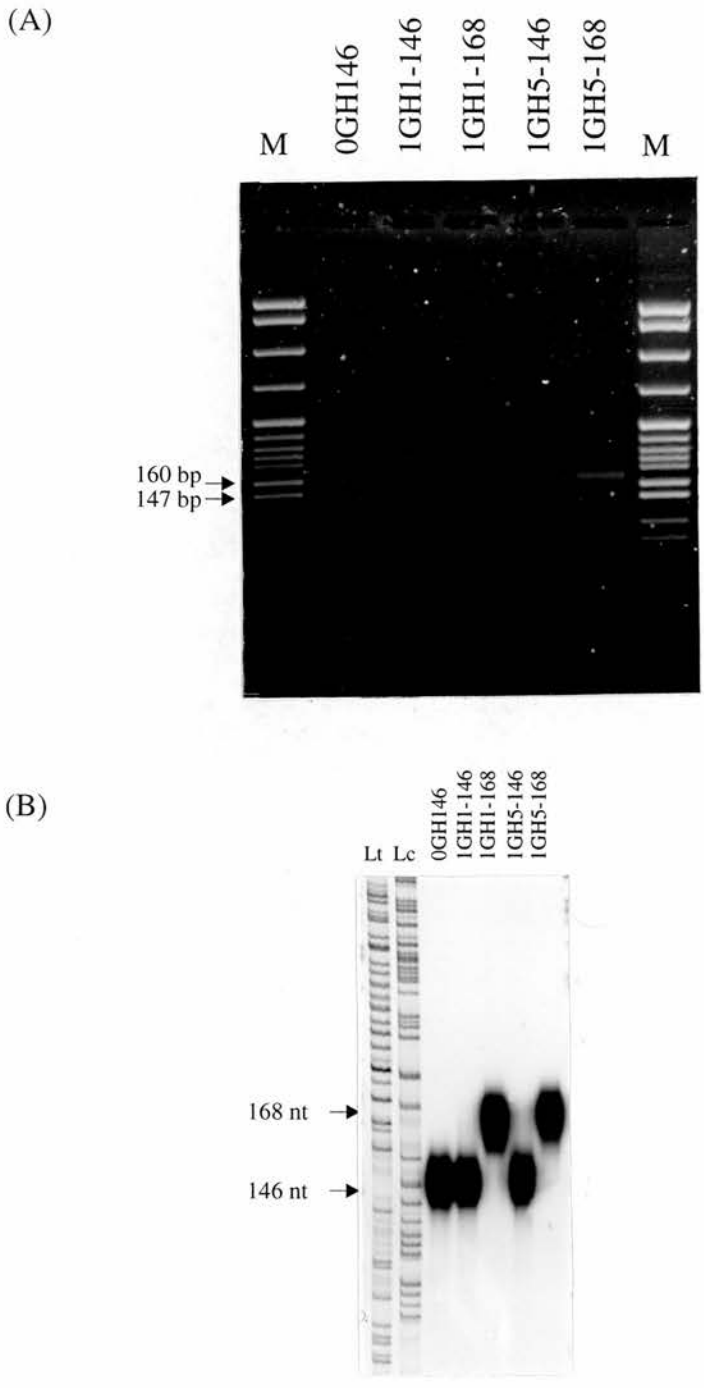


Figure 4-10 (A) A 4.5% metaphor agarose gel analysis of gel-purified core particle and chromatosomal DNAs. (B) A 6% denaturing polyacrylamide gel analysis of 5' end labelled core particle and chromatosomal DNAs. 0GH146: core particle DNA derived from reconstitutes lacking linker histone globular domain. 1GH1-146: core particle DNA derived from GH1-containing reconstitutes. 1GH1-168: chromatosomal DNA derived from GH1-containing reconstitutes. 1GH5-146: core particle DNA derived from GH5-containing reconstitutes. 1GH5-168: chromatosomal DNA derived from GH5-containing reconstitutes. The marker (M) was an MspI digest of pBR322 DNA. Size standards included C and T sequencing reactions of M13mp18 phage DNA (Lc & Lt, respectively).

4.3.2.2 Monomer extension – Mapping plasmids and mapping procedures

Mapping by monomer extension involves annealing labelled DNAs (core particle or chromatosome DNA) to single-stranded mapping plasmids and extending these by the addition of dNTPs and Klenow polymerase. By adding a restriction enzyme in the reaction mix, the nascent double-stranded DNA formed during extension can be cut at a unique location just downstream of (or indeed within) the insert (Figure 4-3). This produces a population of labelled DNA fragments which all share a common 3′ boundary (restriction site) but are distinct in having unique 5′ ends which reflect the location of specific core particle or chromatosome boundaries. The lengths of the fragments produced thus define the location of the core particle or chromatosome boundaries with respect to the restriction enzyme site. In this way, the positions at which the histone octamers or chromatosomes were positioned on the globin DNA can be determined at high resolution.

In this series of experiments, it was important to map both the upstream and downstream boundaries of nucleosomes. For this purpose, two sets of single-stranded mapping clones were produced and employed. These differed with respect to the orientation of their globin insert. For example, XmEcS and EcXmV (Figure 4-2) contain the same 1.5 kb globin insert and provide either the minus single strand (XmEcS) or plus single strand (EcXmV) which permits mapping of either upstream or downstream boundaries, respectively. Phagemids XmEcS, EcSm, LE and Xma were used to define upstream nucleosome boundaries, whereas EcXmV, SmEc, LA and Max were used to define downstream boundaries with respect to the transcription start site of the chicken adult β-globin gene (Figure 4-2).

Products formed during extension reactions were analysed by electrophoresis in 6% denaturing polyacrylamide gels and an example is shown in Figure 4-11. Here, 5′ end-labelled monomers (core particle/chromatosome) prepared from reconstituted chromatin containing core histone alone (0GH), GH1 (1GH1) or GH5 (1GH5) were annealed and extended on single-stranded Max and Xma templates (Figure 4-2) in the presence or absence of the restriction enzyme XbaI (see Figure 4-3 for illustration). In the presence of restriction enzyme (Figure 4-11, lanes 10-14 &

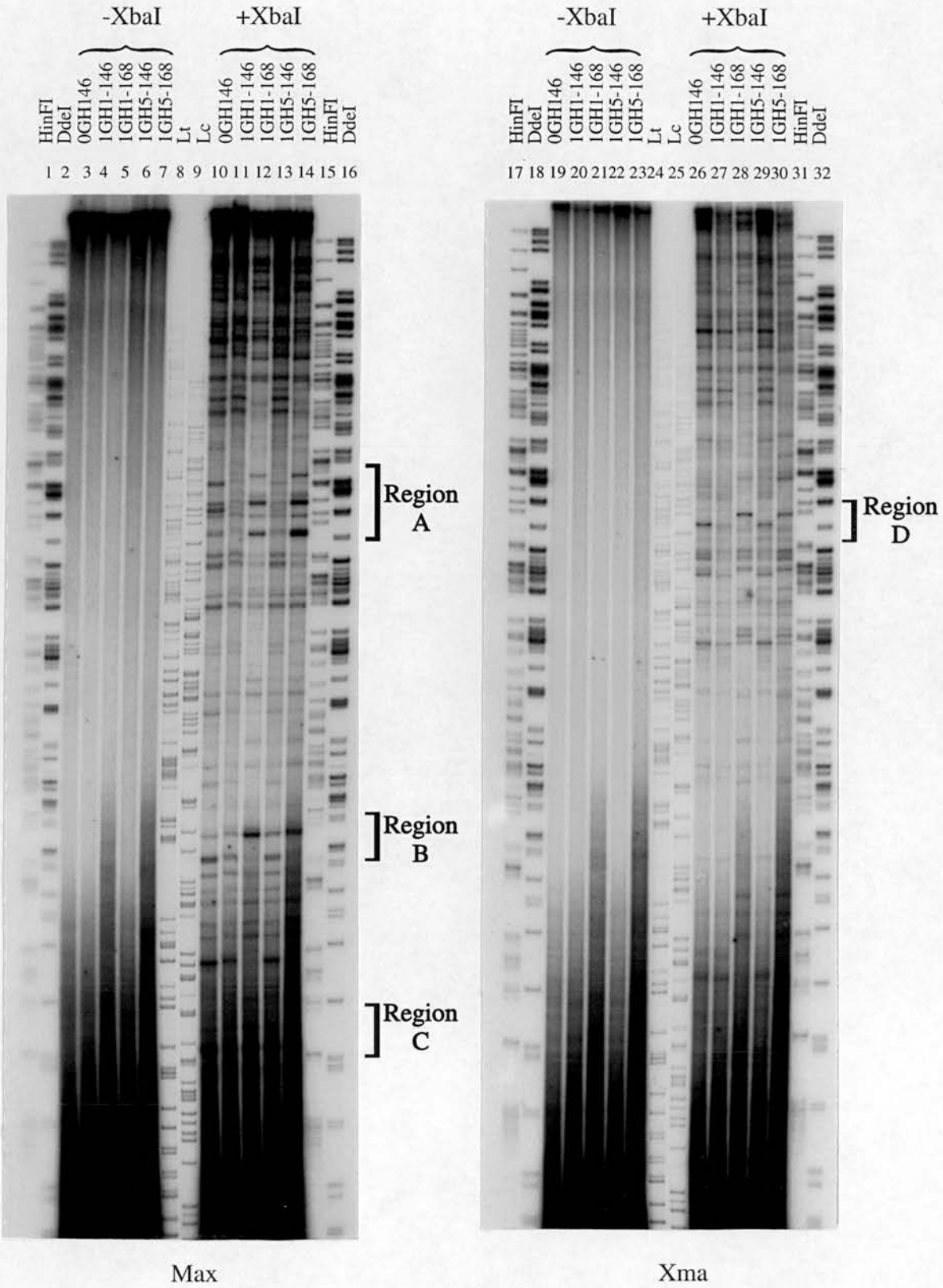


Figure 4-11 6% denaturing polyacrylamide gel analysis of monomer extension products formed on mapping constructs Max and Xma. The names for each sample are as described in Figure 4-10. Size standards included C and T sequencing reactions of M13mp18 phage DNA (Lc & Lt, respectively) and HinfI and DdeI digests of lambda DNA. Specific regions discussed in the text are indicated (A-D).

26-30), extension of the monomer DNA gave rise to a set of discrete, monodisperse bands. The reactions performed in the absence of restriction enzyme (Figure 4-11, lanes 3-7 & 19-23) gave rise to extension products comprising only high molecular weight DNAs and, although there was some background through the lanes, no discrete bands could be detected. The discrete extension products produced in the presence of XbaI are therefore indicative of the boundaries of nucleosome positioning sites.

In order to precisely identify these boundaries, the sizes of the extension products have to be accurately determined. This was achieved by converting densitometry traces obtained from phosphor images of these gels to a linear scale by reference to the various markers run on the gels (Figure 4-11, see Chapter 4.2.11 for detail). By relating fragment sizes to the location of the XbaI site in the mapping plasmid, the location of the nucleosome boundaries with respect to the cap site of the β-globin gene could be determined.

4.3.3 General features of a typical mapping analysis

The monomer extension data arising from analysis and comparison of chromatosome and core particle DNA are characterised by a number of general features. These are illustrated below with reference to Figure 4-11 in which four areas which display examples of these features have been highlighted.

- In Figure 4-11, it is clear that the chromatosome positioning pattern (lanes 12, 14, & 28, 30) is different from the core particle positioning pattern (lanes 10 & 26, respectively). Most of the core particle positioning sites are associated with larger bands corresponding to chromatosome sites. Generally speaking, bands corresponding to chromatosome positioning sites are larger than the corresponding core particle positioning sites.

  In Figure 4-12, Region A from Figure 4-11 has been reproduced and is accompanied by densitometer traces of each of the lanes (10-14). From this

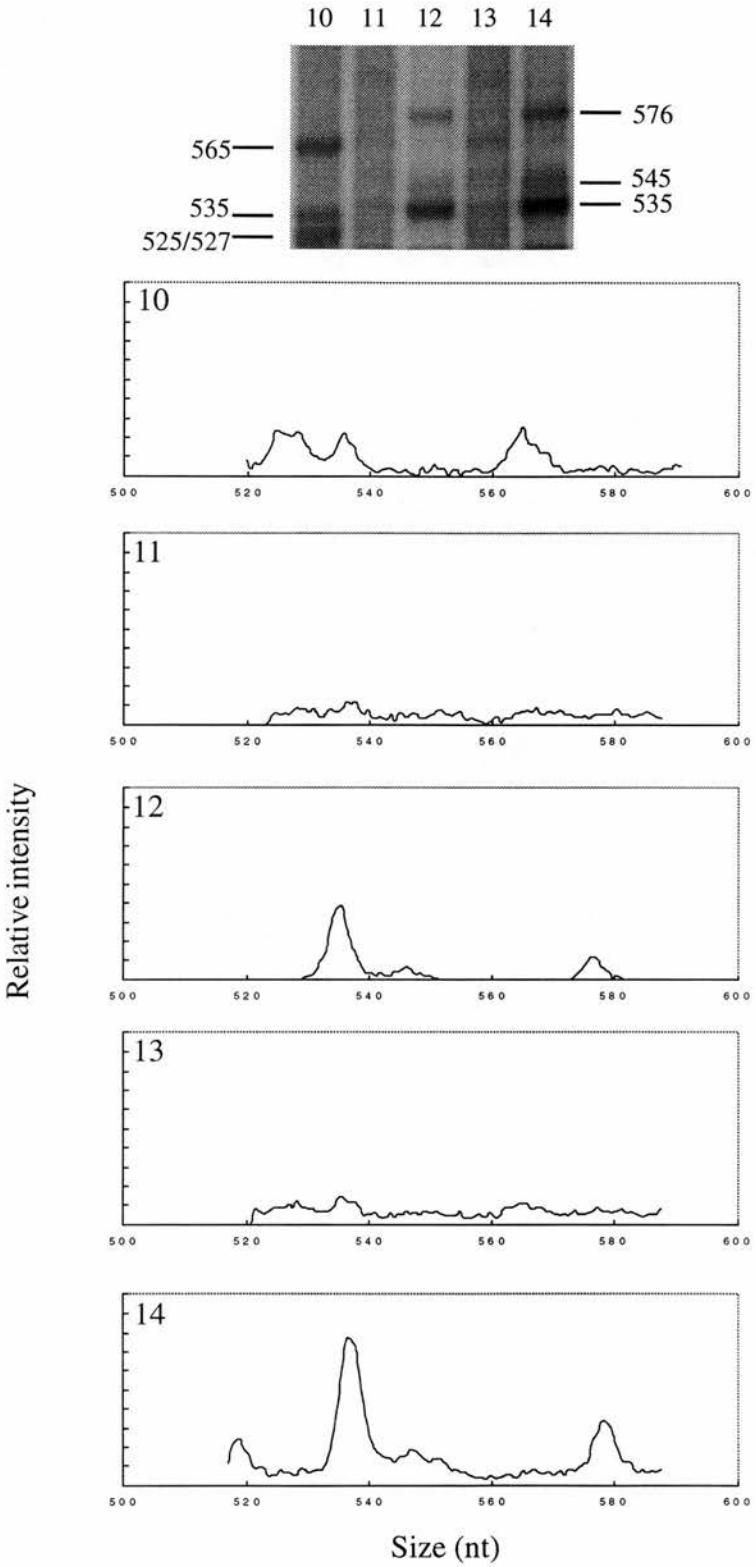


Figure 4-12 Analysis of Region A. Region A from Figure 4-11 is reproduced (top) and the size (nt) of the DNA fragments are indicated. Densitometer traces for each of the lanes 10-14 are presented.

data, the sizes of the various fragments have been determined and are indicated. Figure 4-12 shows that 3 DNA fragments derived from core particles lacking linker histone globular domains (lane 10) have sizes of 565, 535 and 525/527 nucleotides (nt). The DNA fragments of corresponding chromatosome positioning sites derived from GH1 or GH5 reconstitutes (lanes 12 & 14, respectively) have sizes of 576, 545 and 535 nt. In this instance, chromatosome bands are therefore 9 to 11 nt larger than the corresponding core particle bands.

In another example, taken from Region D in Figure 4-11, two DNA fragments derived from core particles lacking linker histone globular domains have sizes of 479 and 496 nt (Figure 4-13; lane 26), whereas the corresponding chromatosome boundaries derived from GH1 or GH5 reconstitutes have sizes of 491 and 511 nt, respectively (Figure 4-13; lanes 28 & 30). This indicates that in this case bands corresponding to chromatosome positioning sites are larger than the corresponding core particle bands by 12 and 15 nt, respectively.

A single strong core particle band from Region C (the other boundary of Region D) has a size of 230 nt (Figure 4-14, lanes 10) whereas the DNA fragments for the corresponding chromatosome boundaries from 1GH1 or 1GH5 have sizes of 230/231 and 234 nt (Figure 4-14, lanes 12 & 14). Here the chromatosome bands are 0 nt and 4 nt larger than the corresponding core particle band.

These results confirm that positioning bands derived from chromatosome reconstitutes are usually larger than the corresponding bands derived from core particle reconstitutes. The difference between the length of a chromatosome band and the length of a corresponding core particle band indicates the length of DNA extending from one side of the core particle and protected from digestion in a chromatosome by the linker histone globular

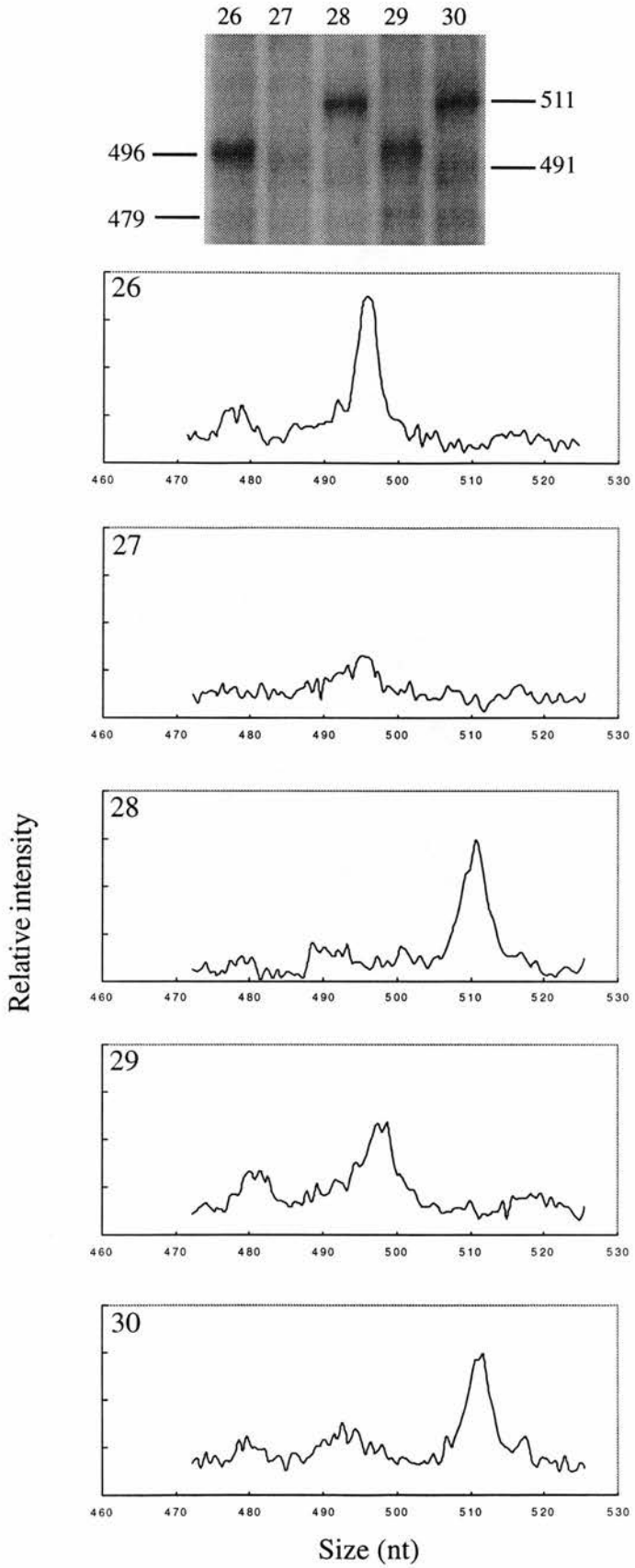


Figure 4-13 Analysis of Region D. Region D from Figure 4-11 is reproduced (top) and the size (nt) of the DNA fragments are indicated. Densitometer traces for each of the lanes 26-30 are presented.

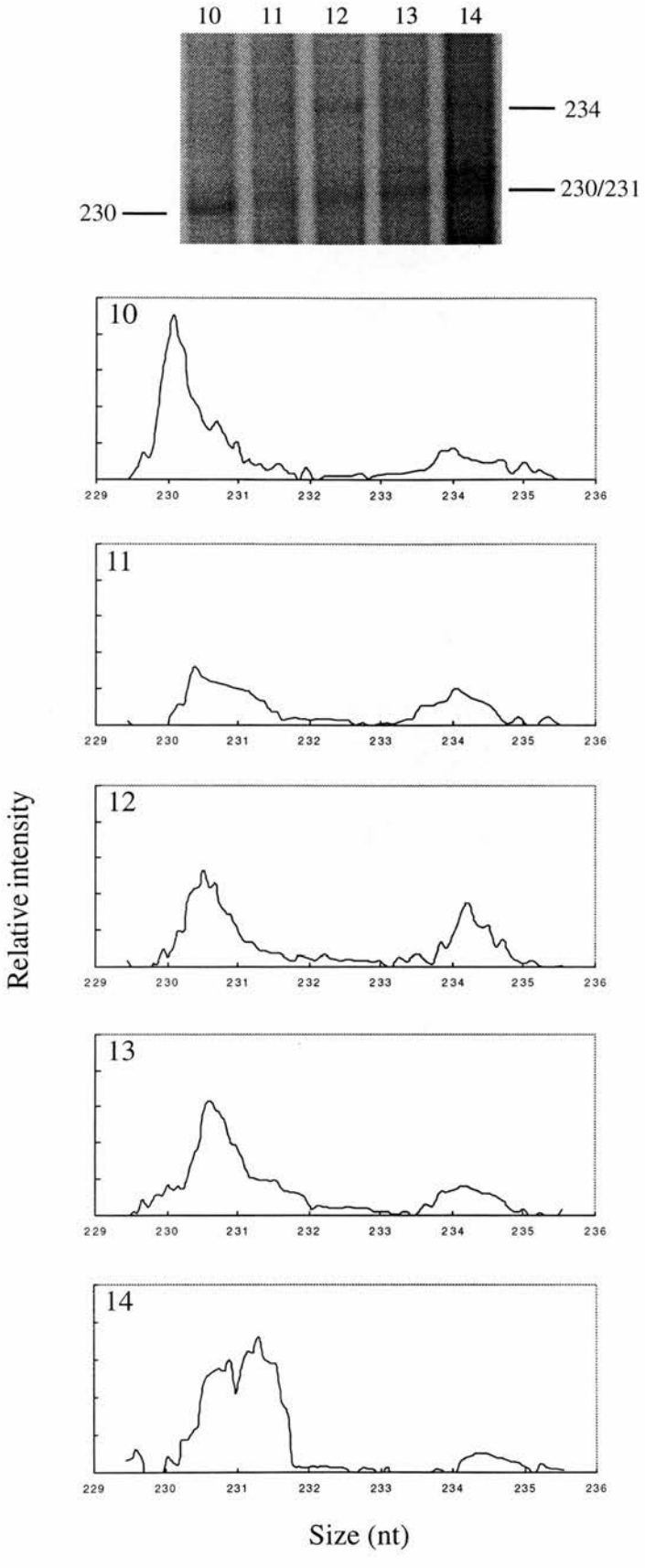


Figure 4-14 Analysis of Region C. Region C from Figure 4-11 is reproduced (top) and the size (nt) of the DNA fragments are indicated. Densitometer traces for each of the lanes 10-14 are presented.

domain. Therefore, these results indicate that the addition of linker histone globular domains can influence DNA extensions to differing degrees.

- Another feature of the raw mapping data (Figure 4-11) is the qualitative similarity in core particle positioning sites and its independence with respect to the source of core particle DNA. The core particle positioning sites derived from reconstitutes lacking linker histone globular domains (Figure 4-11, lanes 10 & 26) and the core particle positioning sites derived from reconstitutes which contained GH1 or GH5 (Figure 4-11, lanes 11, 27 & lanes 13, 29, respectively) are very similar throughout the range of the mapped region. However, there are minor differences in the relative abundance (intensity) of particular sites derived from different chromatins. For example, in Regions A (Figure 4-12) and D (Figure 4-13), the intensities of the bands derived from reconstitutes lacking linker histone globular domains (Figure 4-12, lane 10; Figure 4-13, lane 26) are stronger than the intensities of the core particle bands derived from reconstitutes containing GH1 (Figure 4-12, lane 11; Figure 4-13, lane 27) or GH5 (Figure 4-12, lane 13; Figure 4-13, lane 29). In Region B the intensities of the bands derived from reconstitutes containing GH1 (Figure 4-15, lane 11) or GH5 (Figure 4-15, lane 13) are almost the same as the intensity of the bands derived from reconstitutes lacking linker histone globular domains (Figure 4-15, lane 10).

As the relative intensity of a band is a reflection of fragment abundance in the population, the relative intensity of core particle bands may be an indicator of the stability of the corresponding source chromatosome. The reasoning is that if a chromatosome is particularly stable, it should produce a corresponding core particle more slowly during digestion. Alternatively, if a chromatosome is unstable, it should produce a corresponding core particle at an early stage during digestion. For example, in Region A and D (Figures 4-12 & 4-13, respectively), the chromatosome positioning sites are as strong as the corresponding core particle positioning sites derived from reconstitutes lacking linker histone globular domains. In addition, the core particle bands

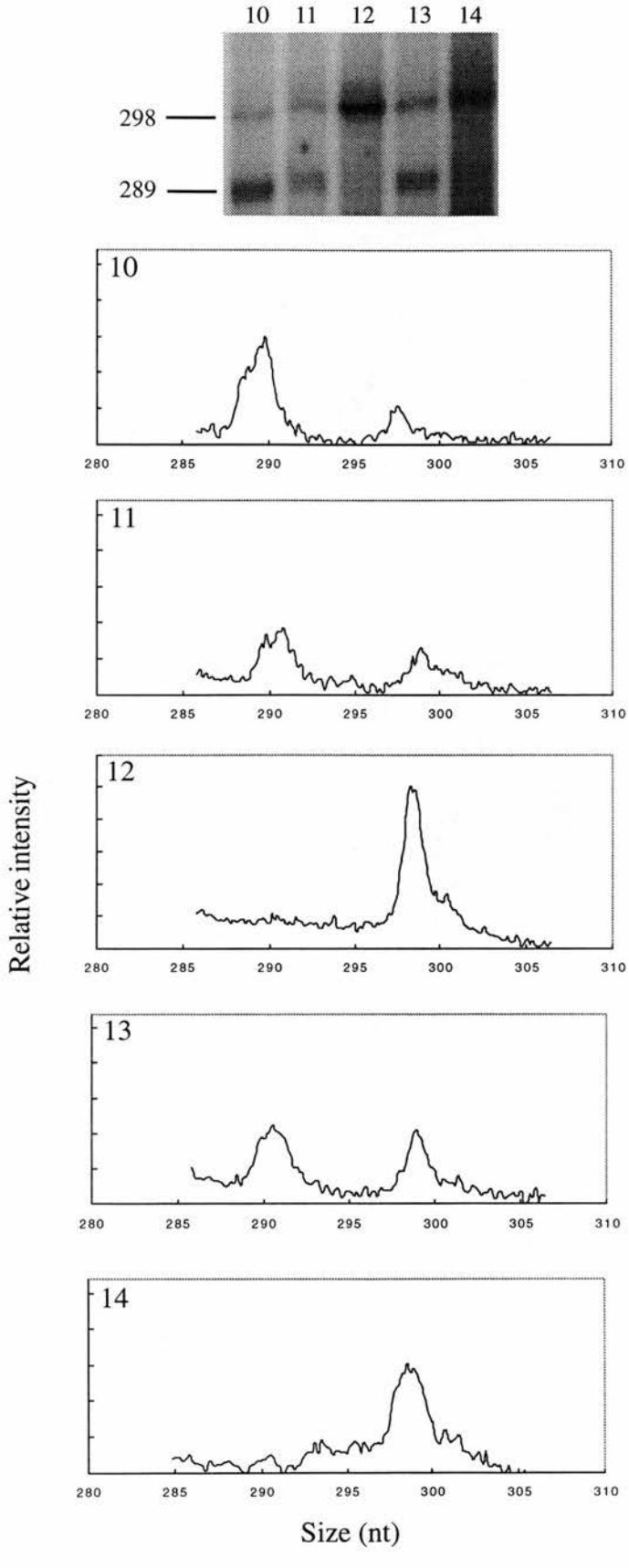


Figure 4-15 Analysis of Region B. Region B from Figure 4-11 is reproduced (top) and the size (nt) of the DNA fragments are indicated. Densitometer traces for each of the lanes 10-14 are presented.

derived from reconstitutes containing GH1 or GH5 are notably weak in intensity (Figure 4-11, lanes 11, 27 & 13, 29; Figure 4-12 & 4-13). In the context of the above proposal, chromatosomes reflected in Regions A and D could be considered to be relatively stable. In contrast, if one considers Region B (Figure 4-15), the intensity of the core particle bands derived from reconstitutes containing GH1 or GH5 (Figure 4-11, lanes 11, 13) are almost as intense as the core particle band derived from the reconstitute lacking linker histone globular domain (Figure 4-11, lanes 10). In this instance, the chromatosome depicted in Region B appears to be relatively unstable.

Although the abundance (frequency of occupation) of core particle positioning sites may differ with respect to different positioning sites and source of chromatin, the location of core particle positioning sites is always the same (Figure 4-11). In other words, core particles derived from chromatosomes occupy the same positions on the DNA as core particles formed in the absence of globular domains. This strongly suggests that the addition of linker histone globular domains does not change nucleosome positioning sites in this system.

- Finally, another notable feature of the analysis is the demonstration that chromatosome positioning sites derived from a GH1-containing reconstitute are very similar to the chromatosome positioning sites derived from a GH5-containing reconstitute (Figure 4-11, lanes 12, 14 & lanes 28, 30); corresponding GH1/GH5 bands are invariably found, although in a few instances, these may differ slightly in abundance (intensity). These observations suggest that the formation of chromatosome positioning sites is independent of the type of linker histone globular domain.

4.3.4 Generation of chromatosome positioning map

By employing all 8 mapping constructs (Figure 4-2), monomer extension was used to map the precise core particle and chromatosome positioning sites, at both their upstream and downstream boundaries, throughout the 1.5 kb region of DNA at the 5` end of the β-globin gene.

A complete gel analysis of the monomer extension experiments, mapping both upstream and downstream boundaries, is shown in Figure 4-16 and Figure 4-17, respectively. As in the previous analysis (Figure 4-11), band sizes were determined after densitometry of phosphor images by reference to markers, and nucleosome positioning sites assigned with respect to the cap site of the β-globin gene.

The general features of the mapping analyses described above are equally applicable to the entire data set. For example, core particle positioning sites derived from the reconstitutes containing GH1 (Figures 4-16 & 4-17, lanes 4, 9) or GH5 (Figures 4-16 & 4-17, lanes 6, 11) are very similar to the core particle positioning sites derived from reconstitutes lacking linker histone globular domains (Figures 4-16 & 4-17, lanes 3, 8), in both upstream and downstream analyses. Again, chromatosome positioning sites (Figures 4-16 & 4-17, lanes 5, 7, 10, 12) are usually larger than the corresponding core particle positioning sites (Figures 4-16 & 4-17, lanes 3, 4, 6, 8, 9, 11), again in both orientations. Finally, chromatosome positioning sites derived from GH1 reconstitutes (Figures 4-16 & 4-17, lanes 5, 10) are similar to the chromatosome positioning sites derived from GH5 reconstitutes (Figures 4-16 & 4-17, lanes 7, 12).

By combining all of the mapping analyses, translational positioning maps for (i) the histone octamer (Figure 4-18B), (ii) GH1- and (iii) GH5-containing chromatosomes (Figure 4-18C & 4-18D, respectively), covering the entire 1.5 kb analysed, can be generated. The core particle positioning map (Figure 4-18B) is very similar to the result of a previous study (Figure 4-18A) (Davey *et al.*, 1995). All the prominent sites determined in this study correspond to strong sites in the Davey

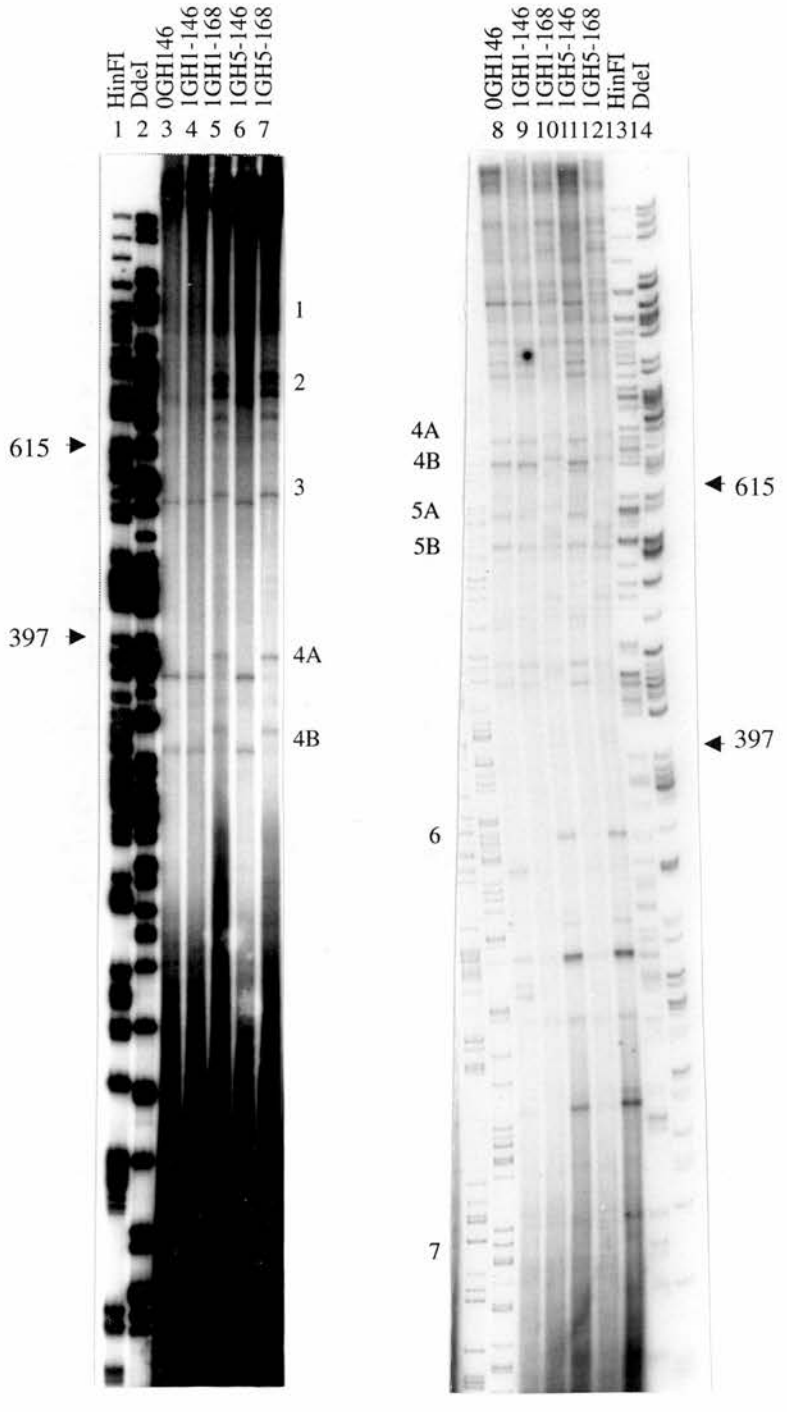


Figure 4-16 6% denaturing polyacrylamide gel analysis of monomer extension products formed on mapping constructs EcSm and LE. The names for each sample are as described in Figure 4-10. Size standards included C and T sequencing reactions of M13mp18 phage DNA (Lc & Lt, respectively) and HinfI and DdeI digests of lambda DNA. Prominent core particle positioning sites previously identified are numbered as in Davey *et al.* (1995) and Davey *et al.* (1997) (5A and 5B distinction) (see Figure 4-2).

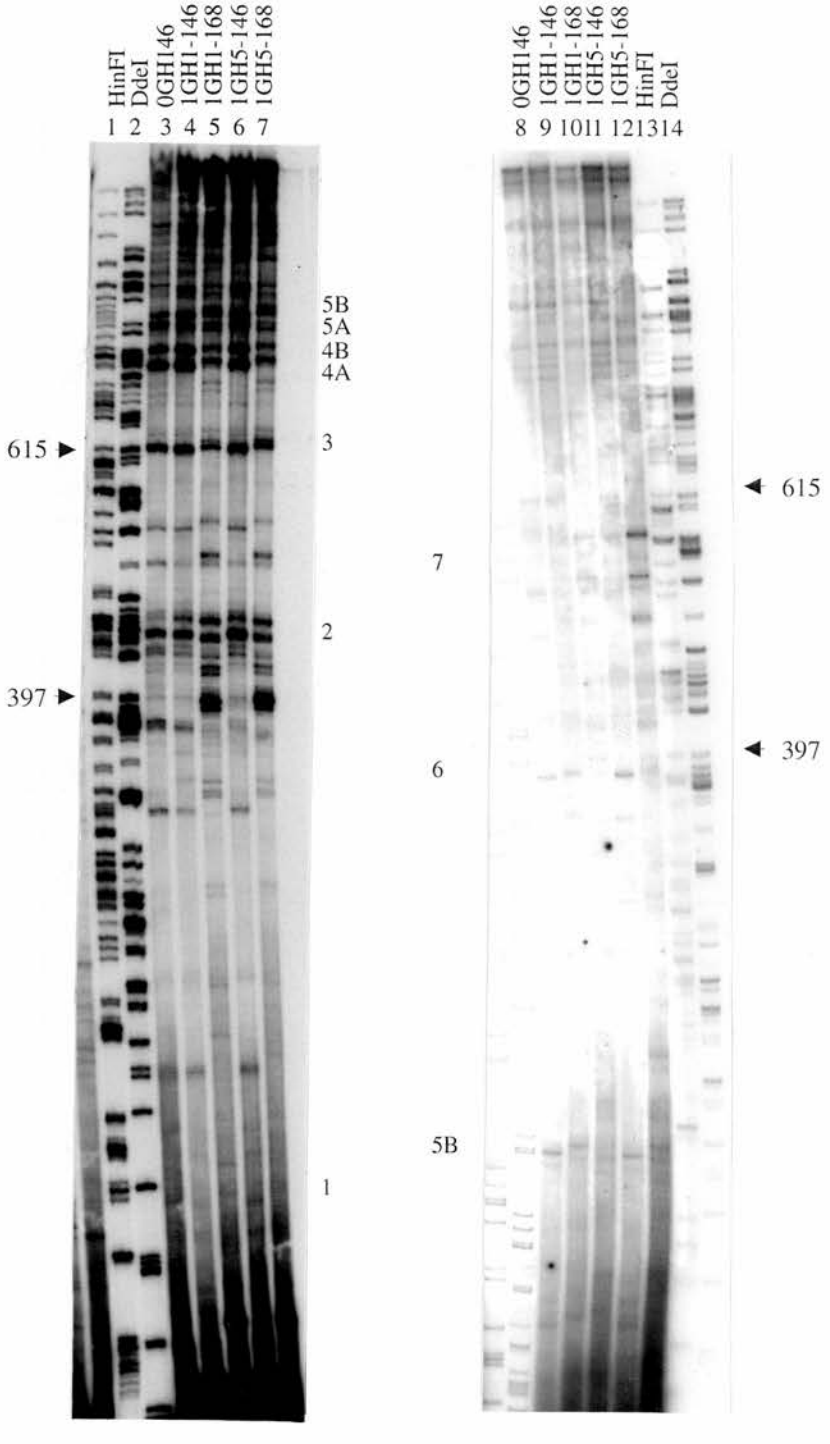


Figure 4-17 6% denaturing polyacrylamide gel analysis of monomer extension products formed on mapping constructs SmEc and LA. Notations as Figure 4-16.

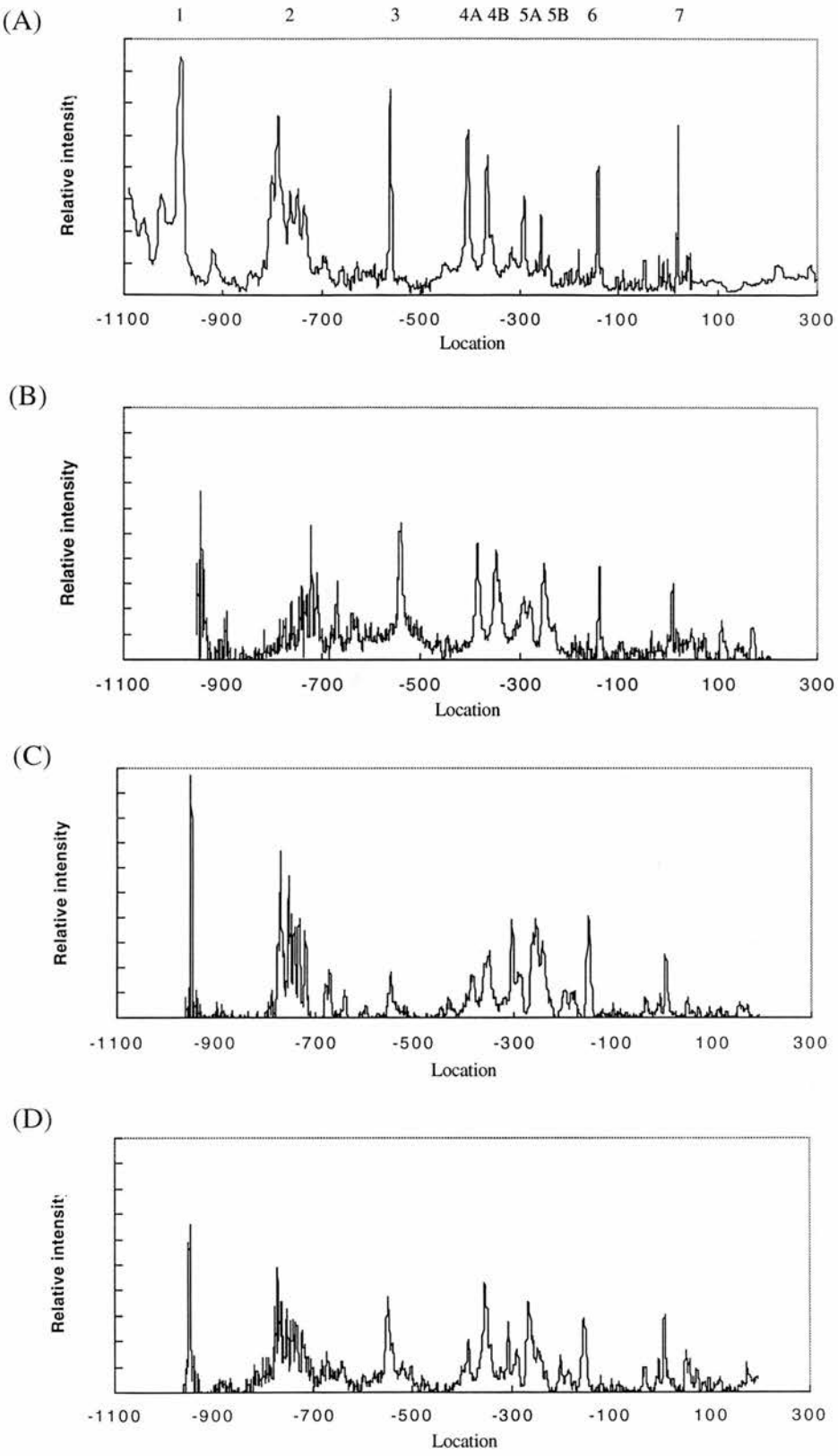


Figure 4-18 Maps of the histone octamer and chromatosome positioning sites for the chicken adult β-globin 5' gene region (1.4 kb). (A) Core histone octamer map taken from Davey *et al.* (1995). (B) Core histone octamer map from this study. (C) Chromatosome positioning map from GH1-containing reconstitutes. (D) Chromatosome positioning map from GH5-containing reconstitutes. The sequence is numbered with respect to the transcription start site of the gene. By assuming a core particle size of 146 bp and a chromatosome size of 168 bp, the maps have been arranged to depict the centres of positioning sites.

study. The only notable exception is the precise position of nucleosome 1 (Figures 4-18A & 4-18B). Davey *et al.* (1995) placed nucleosome 1 at –986 bp (dyad) whereas my studies suggest a location at –952 bp. The previous mapping of nucleosome 1 (Davey *et al.*, 1995) determined only one boundary (only one mapping construct was used). Furthermore, the size of the fragment which determined the location of nucleosome 1 was large and migrated towards the top of the sequencing gel. There is likely to be substantial inaccuracy in measuring large fragment lengths (> 800 bp) in a sequencing gel. In this study the location of nucleosome 1 was determined on 4 different mapping plasmids at both boundaries and, therefore, the location of nucleosome 1 may have been measured more precisely.

The chromatosome positioning sites corresponding to all of the strong core particle positioning sites previously determined by Davey *et al.* (1995) have been analysed in detail. Below, the properties relating to positioning at sites 3, 4A, 5B and 7, are discussed in depth (Figure 4-18).

Positioning at sites 3 and 4A: Nucleosomes 3 and 4A are well defined in both the upstream and downstream analyses. In Figure 4-19 these data are reproduced and supplemented by densitometry traces to facilitate analysis.

In Figure 4-19A, the upstream boundaries of core histone octamer positioning sites 3 and 4A map at –633 and -485 (relative to the β–globin cap site). Core particle positioning sites derived from reconstitutes containing GH1 (1GH1-146; Figure 4-19, lane 4) or GH5 (1GH5-146; Figure 4-19, lane 6) have the same upstream coordinates as the core particle positioning sites derived from reconstitutes lacking linker histone globular domains (0GH-146; Figure 4-19, lane 3).

Figure 4-19B demonstrates that chromatosome positioning at sites 3 and 4A is independent of the type of globular domain employed during reconstitution and the upstream boundaries determined by these analyses are located at –644 and –493.

In Figure 4-19C, a similar analysis has been employed to determine the

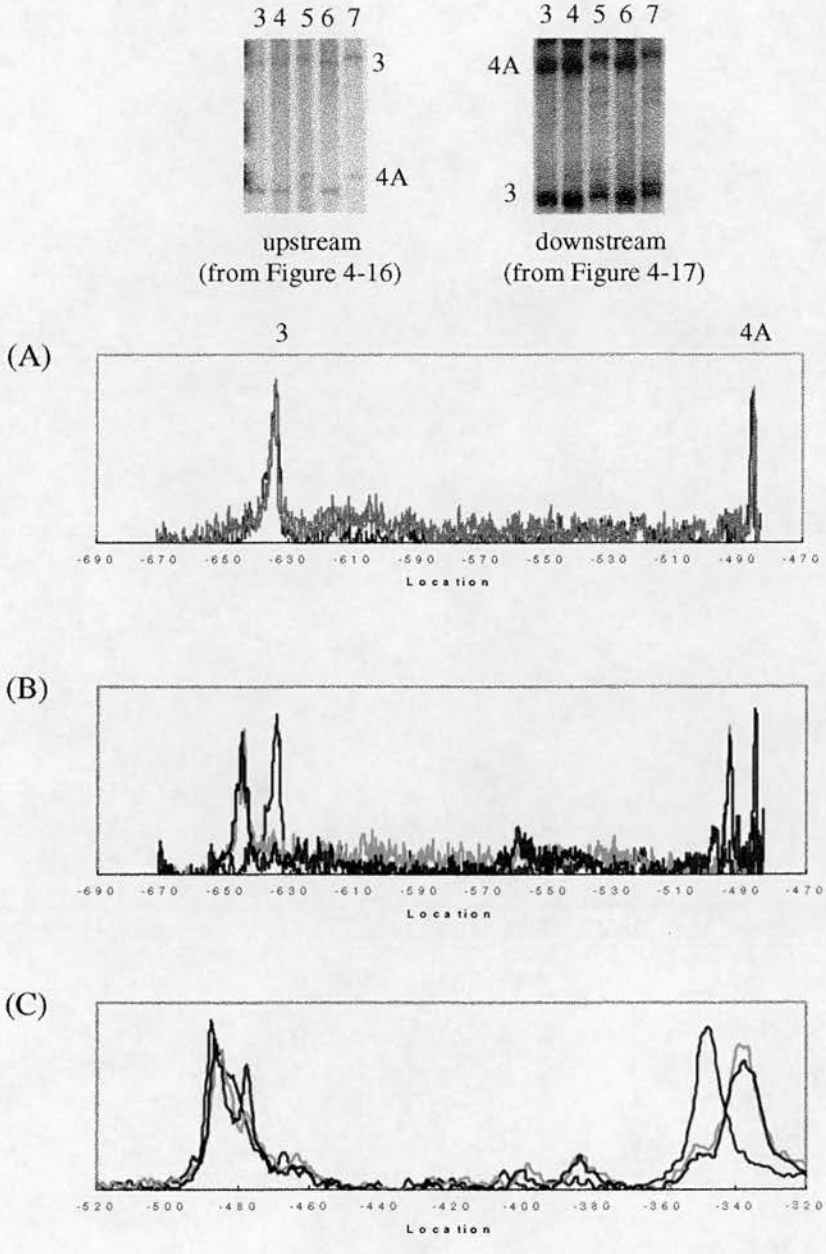


Figure 4-19 An example of the determination and comparison for nucleosomes 3 and 4A. Relevant regions of the gel analyses (Figures 4-16 & 4-17) are reproduced and lane numbering is as in Figure 4-16 & 4-17. (A) Densitometry traces indicating upstream boundaries for core particle derived from 0GH146 (blue), 1GH1-146 (brown) and 1GH5-146 (orange). (B) Densitometry traces indicating upstream boundaries for core particle derived from 0GH146 (blue) and for chromatosome derived from 1GH1-168 (pink) and 1GH5-168 (green). (C) Densitometry traces indicating downstream boundaries for core particle derived from 0GH146 (blue) and for chromatosome derived from 1GH1-168 (pink) and 1GH5-168 (green). Location is with respect the β-globin gene cap site.

downstream boundaries for 3 and 4A cores and chromatosomes. Core boundaries map to position –487 (3) and –347 (4A) whereas chromatosome boundaries map to –477 (3) and -337 (4A).

This analysis demonstrates that for site 3 an extra 11 bp was protected at the upstream boundary in the chromatosome whereas an extra 10 bp of DNA was protected at the downstream boundary. Similarly, for site 4A an extra 8 bp was protected at the upstream boundary and an extra 10 bp was protected at the downstream boundary. Thus, for nucleosomes 3 and 4A the DNA extension in the chromatosome appears to be approximately equally distributed at each end of the core particle.

Positioning at site 5B: The data presented for the analysis of core histone octamer positioning site 5B are described in the same manner as above. In Figure 4-20A, the upstream boundary of the core histone octamer positioning site maps at –325 and the corresponding chromatosomes appear to have upstream boundaries located at –325, -335 and –345. The result is independent of the type of globular domain employed during reconstitution.

In Figure 4-20B, a similar analysis has been employed to determine downstream boundaries for the 5B core particle and chromatosome positioning sites. These were mapped to –178 bp (core) and to –178 and -169 bp (chromatosome). Again their location was independent of globular domain type.

This analysis demonstrates that for site 5B the extra DNA protected at the upstream boundary in a chromatosome appears to be 0, 10 or 20 bp. For the downstream boundary, an extra 0 and 9 bp were protected in the chromatosome. Thus, for Nucleosome 5B the additional DNA in the chromatosome appears to be asymmetrically (20+0) and symmetrically (10+10) distributed at each end of the core particle.

Positioning at site 7: In Figure 4-11, the site 7 upstream and downstream boundaries

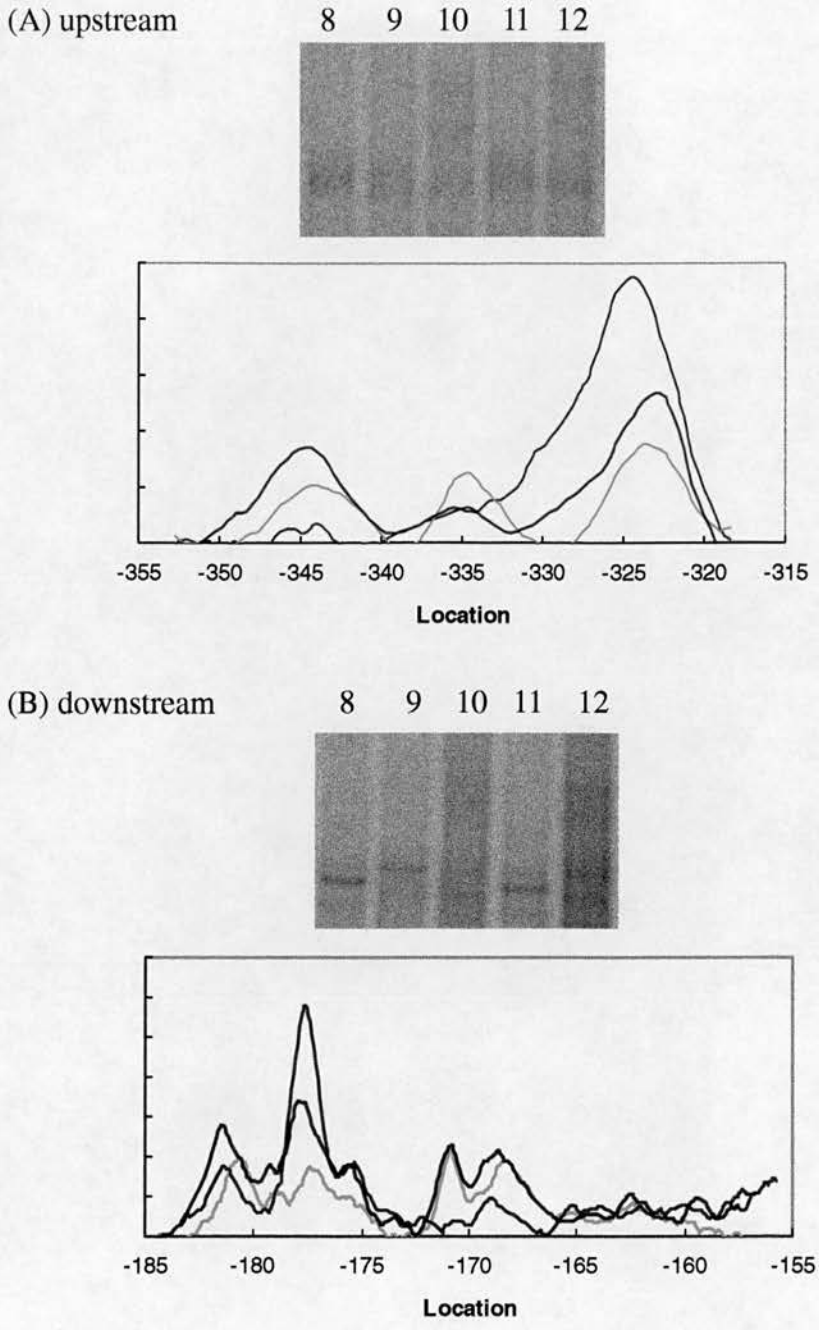


Figure 4-20 An example of the determination and comparison for nucleosomes 5B. Relevant regions of the gel analyses (Figures 4-16 & 4-17) are reproduced and lane numbering is as in Figures 4-16 & 4-17. (A) Densitometry traces indicating upstream boundaries for core particle derived from 0GH146 (blue), 1GH1-146 (brown) and 1GH5-146 (orange). (B) Densitometry traces indicating upstream boundaries for core particle derived from 0GH146 (blue) and for chromatosome derived from 1GH1-168 (pink) and 1GH5-168 (green). (C) Densitometry traces indicating downstream boundaries for core particle derived from 0GH146 (blue) and for chromatosome derived from 1GH1-168 (pink) and 1GH5-168 (green). Location is with respect the β-globin gene cap site.

are depicted in Regions D and C, respectively. Figure 4-13 (expansion of Region D) shows a 15 nt extension for the upstream chromatosome boundary at site 7 and Figure 4-14 (expansion of Region C) shows that 4 nt and 0 nt extensions are found at the downstream boundary. Thus, in this case, the additional DNA in the chromatosome appears to be asymmetrically distributed (15+4)at the ends of the core particle.

4.3.4.1 DNA extension lengths in chromatosomes

By applying the type of analysis outlined above, it was possible to compare the DNA extensions in chromatosomes, relative to core particles, for all the major positioning sites throughout the 1.5 kb region mapped (Figure 4-2). A summary of this analysis is shown in Table 4-1. The results indicate that DNA extensions for GH1-containing chromatosomes were the same as the DNA extension for GH5-containing chromatosomes. It also illustrates that in most instances, the additional 20 bp of DNA in a chromatosome is made up of an extra ~10 bp at each boundary of the core particle. There are exceptions to this rule such as nucleosomes 5B and 7. For nucleosome 5B, protected extensions fall into two types, 10+10 bp and 20+0 bp. For nucleosome 7, there is one form of protection, 5+15 bp.

From the above results, a general classification of chromatosome protection for the major nucleosome positioning sites has been established. The DNA extensions did not show exact 5+15, 10+10 or 20+0 bp pairings. Instead, variable lengths of DNA extensions were exhibited. To characterise this point further, 37 nucleosomes were studied in detail to determine their extension lengths. These nucleosomes were selected from regions spread over the 1.5 kb mapping region and included the data from the 9 major nucleosomes detailed above (Table 4-1).

74 different DNA extension lengths were obtained for the 37 nucleosomes selected. Figure 4-21A shows the distribution of DNA extension length. The graph illustrates that there is a strong tendency toward a distribution centred about 10 bp. The data are bell-shaped and its mean and median fall at the centre of symmetry,

Positioning sites	Reconstitution condition			
	4°C		37°C	
	upstream	downstream	upstream	downstream
1	10	10	8	9
2	13	12	10	8
3	12	9	8	11
4A	13	10	11	9
4B	10	11	9	9
5A	12	13	10	10
5B	20/10	0/ 9	19/10	0/10
6	12	9	11	8
7	15	4	14	6

Table 4-1 Summary of the lengths of DNA extensions associated with chromatosomes compared to core particles.

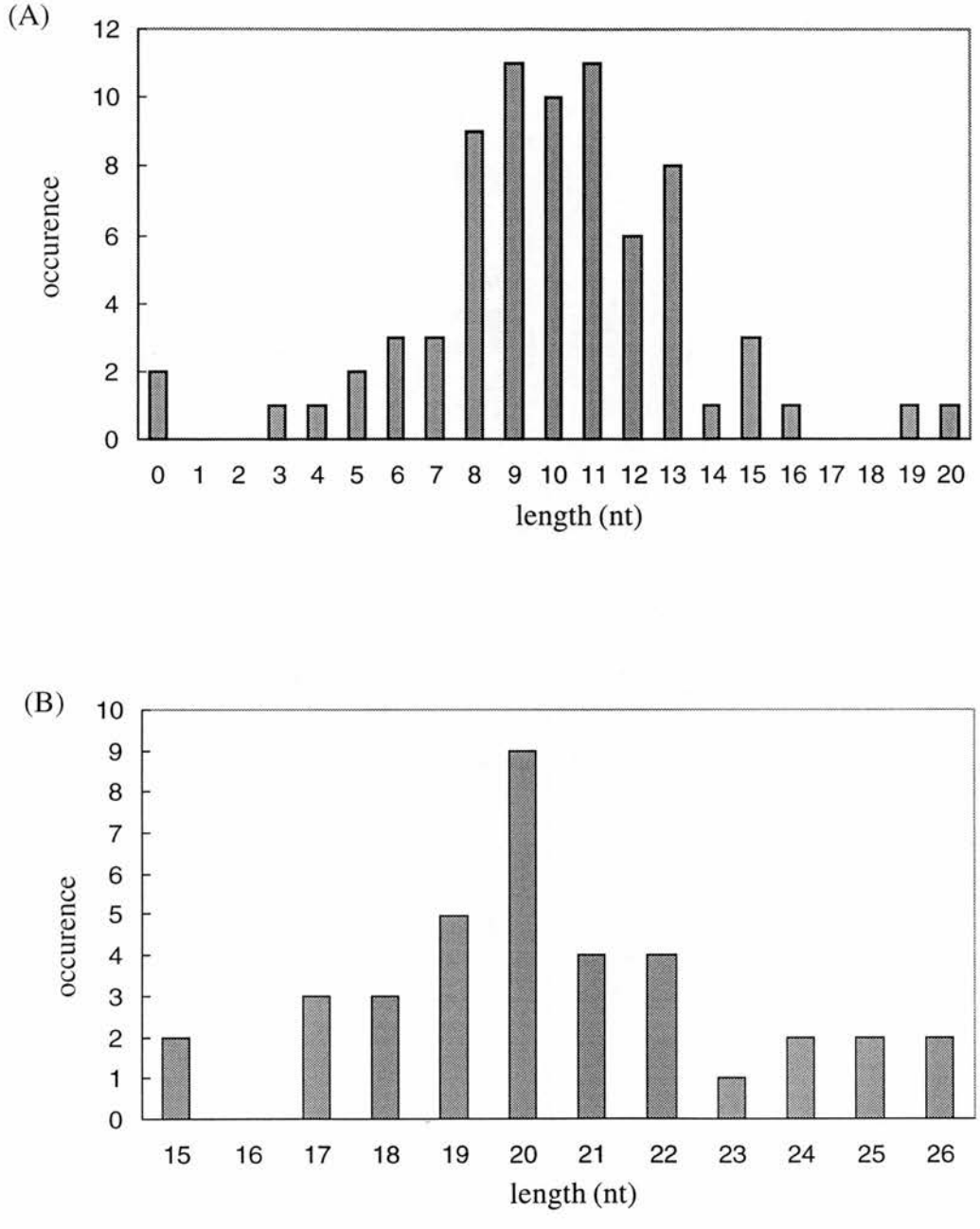


Figure 4-21 Distribution of lengths of extra DNA associated with chromatosomes compared to core particles. (A) Distribution for 74 different DNA extension lengths obtained from 37 selected nucleosomes. (B) Distribution for the sums of paired lengths of DNA extensions for each of the 37 nucleosomes. List of coordinates (position of dyad axis) of all 37 selected nucleosomes: -952, -931, -829, -797, -787, -784, -769, -763, -750, -746, -737, -730, -721, -696, -660, -626, -560, -404, -365, -336, -320, -306, -291, -284, -255, -235, -224, -173, -163, -140, -130, -65, -51, -39, -2, +13, +22.

indicating a normal distribution. The distribution provides an average DNA extension of 10.4 ± 0.40 bp and a 99% confidence interval (μ) of 10.4 ± 1.05 bp.

By summing the paired lengths of DNA extensions for each of the same 37 nuclesomes, the total length of additional DNA associated with a chromatosome (compared to a core particle) can be obtained. This analysis (Figure 4-21B) indicates a strong tendency toward a distribution centred about 20 bp. The distribution indicates that the average sum of paired DNA extensions is 20.8 ± 0.45 bp and the 99% confidence interval (μ) is 20.8 ± 1.21 bp.

The above analysis indicates that the DNA which extends from a core particle to form a chromatosome amounts to about 21 bp and is almost commonly equally distributed in 10 bp units at both ends of the core particle.

4.3.4.2 Quantitative aspects of chromatosome positioning

It has already been mentioned that quantitative information concerning the relative stability of different chromatosomes may be derived from the positioning analyses. An example of this property is seen by examining the cluster of overlapping positioning sequences at site 2 (Figures 4-16 & 4-17). In this (~250 bp) region, there are 12-14 overlapping core histone octamer positioning sites. Most of these sites can be directly associated with a corresponding chromatosome site. To illustrate this for both the upstream and downstream analyses, densitometry traces for core particle and chromatosome positioning have been aligned in Figure 4-22.

If one considers the intensities of the bands corresponding to core/chromatosome sites 2^a, 2^b, 2^c, 2^h, 2^i, 2^j, 2^k, 2^l, 2^m and 2^n (Figure 4-22), one observes that their intensities do no change substantially in response to the addition of GH1. In other words, corresponding core particles and chromatosomes have equivalent intensities. However, the intensities of bands corresponding to core/chromatosome sites 2^d, 2^e, 2^f and 2^g are significantly enhanced in response to the addition of GH1, at both upstream and downstream boundaries. In other words,

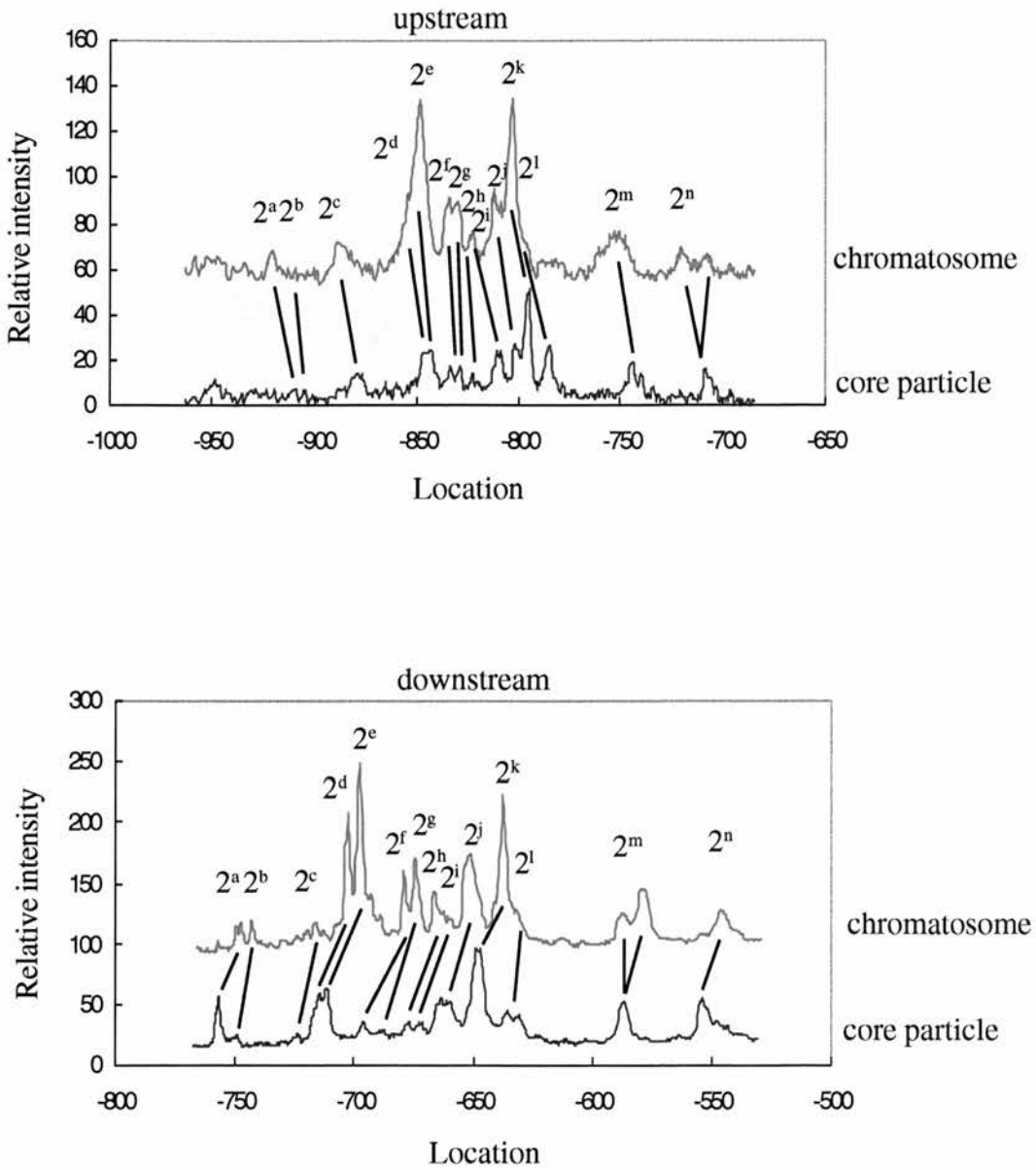


Figure 4-22 Densitometry traces of chromatosome and core particle boundaries in the nucleosome 2 clusters. The upstream boundary data were derived from analyses not shown here (however, see Figure 4-16). The downstream boundary data were derived from lanes 3 (core particle) and 5 (chromatosome) in Figure 4-17. The lines join corresponding core particle (blue) and chromatosome (red) bands. Peaks have been labelled from 2^a to 2^n for the purposes of discussion.

chromatosomes are more strongly representative than their corresponding core particle. This is especially notable for positioning site 2^e (Figure 4-22). Both GH1- and GH5-containing reconstitutes display the same enhancement feature (Figure 4-17).

The selective enhancement of particular positioning sites seen when one compares chromatosome and core particle boundaries could be explained in two ways. Firstly, it may indicate that GH1/GH5 has a particular preference to bind to a specific nucleosome and thereby enhance its protection. The alternative explanation is that specific chromatosomes are simply more resistant to nuclease digestion and accumulate in the population during digestion. Of course, these explanations may be related in that preferential binding may confer greater resistance.

In order to further investigate the relative protection of particular chromatosomes, an experiment was undertaken to assess the stability of chromatosomes as the function of MNase digestion. Aliquots of core histone reconstituted chromatin were titrated with one molecule of GH1/GH5 per core histone octamer. These chromatins were digested with MNase under standard conditions but with different trimming times. The chromatosomal DNAs were purified from an agarose gel and were examined in a 4.5% agarose gel (Figure 4-23A) and, after labelling, in a 6 % denaturing polyacrylamide gel (Figure 4-23B). The purified chromatosomal DNAs were then employed for monomer extension.

The results of this analysis are shown in Figure 4-23C. Once again the data demonstrate that GH1- and GH5-containing chromatosomes (lanes 1 to 4, lanes 5 to 8, respectively) are very similar throughout the mapping region. However, within either analysis some features of the chromatosome binding pattern vary as a function of trimming time. Figure 4-23C shows that some chromatosome bands retain a constant intensity throughout the trimming period (2^a, 2^b, 2^j, 2^k, 2^l, 2^n, 4A and 4B). Some chromatosomes display enhancement of intensity as a function of increasing trimming time (2^c, 2^d, 2^e, 2^f, 2^g, 2^h and 2^i). However, there are some notable exceptions which display properties suggestive of chromatosome degradation (2^m

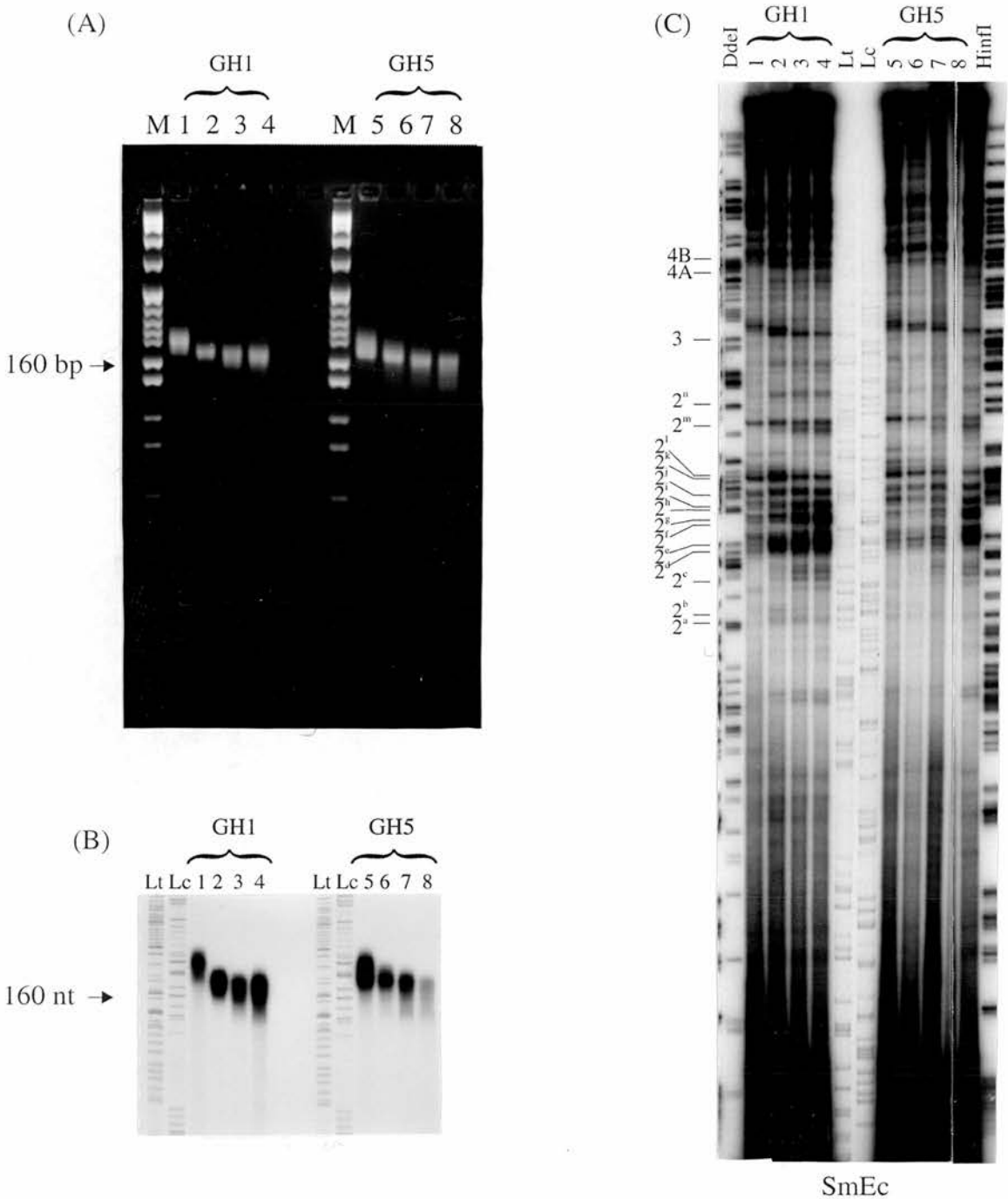


Figure 4-23 Stability of chromatosomes as a function of MNase digestion. (A) 4.5% metaphor agarose gel analysis of purified, 5' end labelled, chromatosomal DNA. (B) A 6% denaturing polyacrylamide gel analysis of chromatosomal DNA. (C) 6% denaturing polyacrylamide gel analysis of monomer extension products derived from GH1- (lanes 1 to 4) or GH5- (lanes 5 to 8) containing reconstitutes. Selected products are indicated in accordance with Figure 4-22. Trimming times were 35 seconds (lanes 1 & 5), 70 seconds (lanes 2 & 6), 105 seconds (lanes 3 & 7) and 140 seconds (lanes 4 & 8). Size standards are as described in Figure 4-11.

and 3).

To analyse these points further, densitometry traces of the relevant regions (Figure 4-23, lanes 1 to 4) are presented in Figure 4-24. It can be seen that for nucleosome 3 there is initially a strong band which maps to –478 and corresponds to the chromatosome boundary for this nucleosome. As digestion proceeds this band rapidly decreases and is almost entirely replaced by a shorter fragment which maps to –488 and corresponds to the core particle boundary for this nucleosome. This behaviour suggest that chromatosome 3 is being degraded during digestion to a core particle-like fragment.

Similarly, for Nucleosome 2^m, there is initially one strong band which maps to –600 and corresponds to the chromatosome boundary for this nucleosome. Although this band retains the same intensity throughout the trimming duration, another lower band, which maps to –609 and corresponds to the core particle boundary for this nucleosome, appears gradually as digestion proceeds (Figure 4-24). This is an example of lesser chromatosome instability

The scheme outlined in Figure 4-25 is an attempt to explain the above observations through chromatosome sensitivity to nuclease digestion. A chromatosome which is not sensitive to MNase (a stable chromatosome) will produce a full length 168 nt monomer fragment. These monomers when applied to monomer extension will indicate a normal chromatosome boundary, 10 nt larger than the corresponding core particle boundary. On the other hand, a chromatosome which is sensitive to internal single-stranded nicking by MNase (an unstable chromatosome) could produce 158 and 10 nt monomers which may co-fractionate with full length chromatosome DNA. When used for monomer extension the 158 nt fragment will indicate a "core" boundary. The 10 nt fragment may be too short for mapping by monomer extension or alternatively may be lost during purification of the labelled fragment.

During the course of analysis, it was noted that for some nucleosomes which

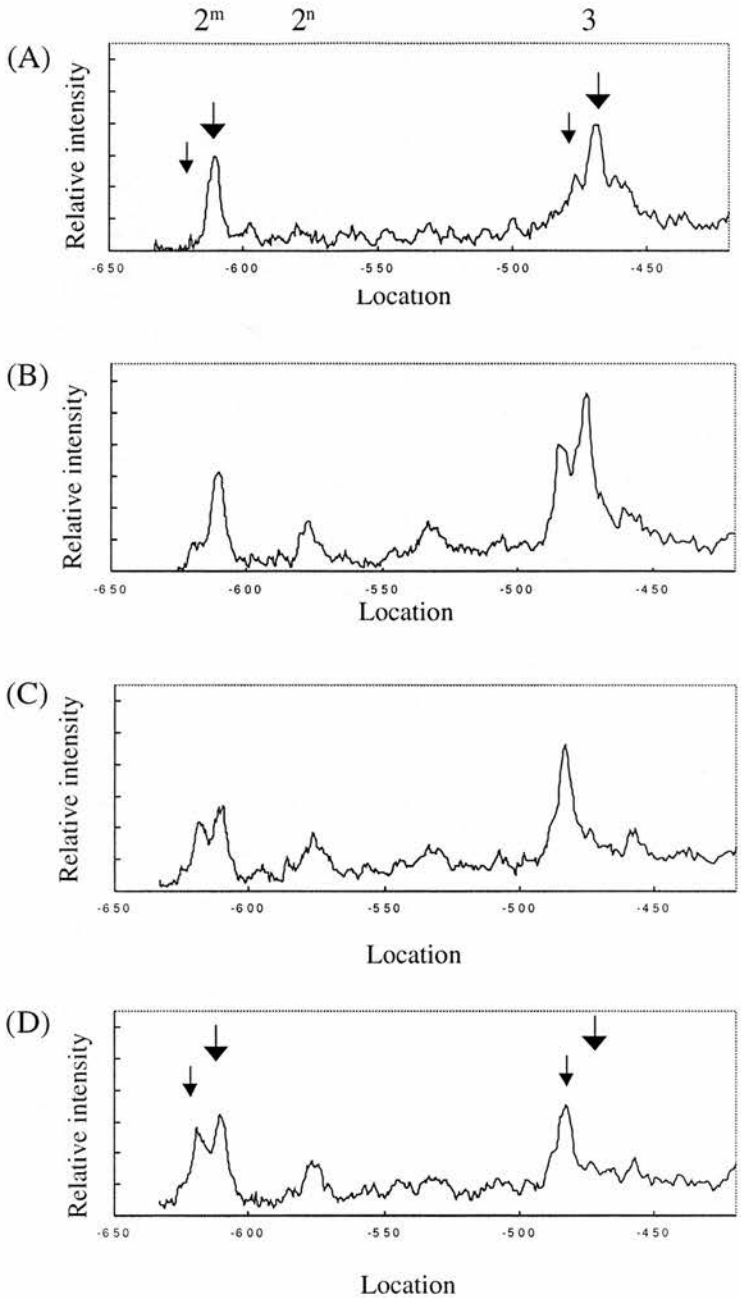


Figure 4-24 Changing patterns of chromatosome protection for selected nucleosomes as a function of MNase trimming time(A) 35 sec (B) 70 sec (C) 105 sec and (D) 140 sec. Chromatosome positioning boundries are indicated by the large arrow and predicted/expected core particle boundaries are indicated by small arrows. Location is with respect to β-globin gene cap site.

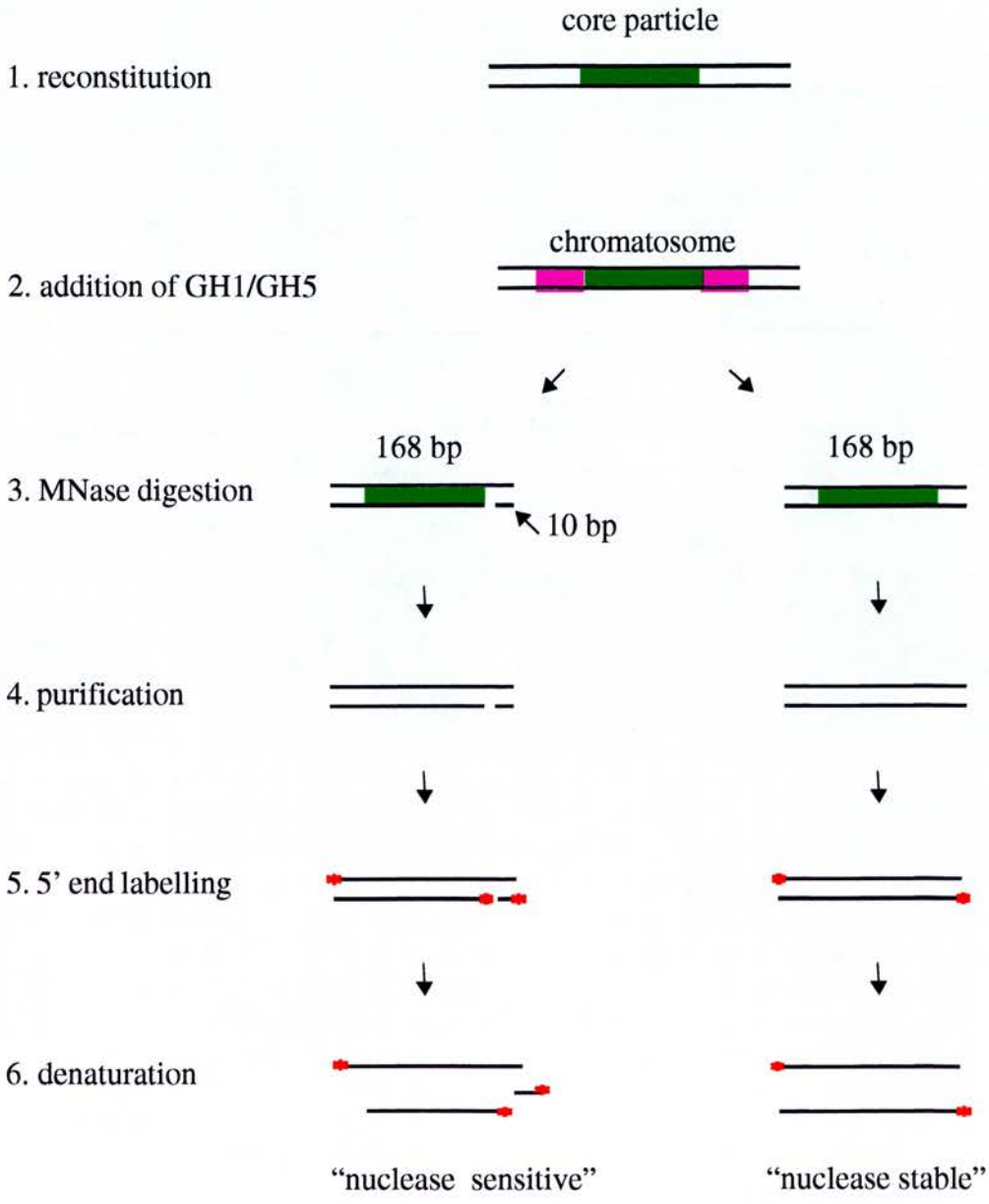


Figure 4-25 Possible scenario for the introduction, by MNase, of single-strand nicks into chromatosome DNA as discussed in the text.

displayed a 0 bp extension at one end, there was not always a corresponding 20 bp extension at the other end. In the light of the above explanation, these chromatosomal DNAs could be sensitive to MNase and subject to single-stranded nicking during their preparation.

In summary, these results indicate (i) that the affinity of GH1/GH5 for different chromatosome positioning sites may be variable, (ii) that there is no difference between the positioning maps for GH1-containing chromatosomes and GH5-containing chromatosomes, (iii) that the extra 20 bp protected by linker histone globular domains are usually, but not always, symmetrically disposed and (iv) that different chromatosomes display variability in their sensitivity to MNase. This latter effect may be inherent to the DNA at the boundaries of the chromatosome.

4.3.5 Effect of different reconstitution conditions on nucleosome positioning

Positioned nucleosomes have been proposed to demonstrate a dynamic behaviour, which effects a constant shuffling between available positioning sites (Pennings *et al.*, 1991; Meersseman *et al.*, 1992; Pennings *et al.*, 1994). Pennings *et al.* (1991) found that histone octamers located at the cluster of positioning sites on sea urchin 5S rDNA were in dynamic equilibrium and that redistribution of octamers between available positions occurred at low ionic strength (10 mM Na^+) at 37 °C but not at 4°C. They therefore suggested that histone octamer mobility was temperature-dependent. Furthermore, subsequent experiments have demonstrated that this temperature-dependent dynamic behaviour could also apply to bulk mononucleosomes and nucleosomes reconstituted onto sequences of the Alu family of ubiquitous repeats (Meersseman *et al.*, 1992). Although the movement of nucleosomes increases with increasing temperature, there are boundaries to this mobility. The mobility of histone octamers was limited to (i) the cluster of positions around the dominant sequence and (ii) the positioning sites were related by having the same rotational setting. Bradbury and his co-worker proposed that short range sliding was a general phenomenon that was dependent on the underlying DNA

sequence and its position relative to the histone octamer (Meersseman *et al.*, 1992; Pennings *et al.*, 1994).

The results presented above demonstrate that the addition of linker histone globular domains to reconstituted chromatin does not appear to alter nucleosome positioning. Instead, chromatosome positioning sites appear to be determined by the core particle during the reconstitution process before the addition of linker histone globular domains. Almost all of the chromatosome positions which have been mapped can be accounted for on the basis of the established core positions. There is little evidence to suggest that new positions are being determined by virtue of the presence of linker histone globular domains. As the studies described above were carried out by addition of GH1/GH5 in low salt (80 mM Na^+) at 4°C, it was considered important to determine the extent to which linker histone globular domains could influence or determine nucleosome positioning if they were added during the process of reconstitution and under conditions thought to be more favourable to octamer mobility (dynamic condition).

Again monomer extension was employed to study nucleosome positioning in a "dynamic" reconstitution system. The strategy was the same as the standard monomer extension procedure except that linker histone globular domains were added during reconstitution and subsequent incubation was carried out at 37°C. Briefly, reconstitution of core histones, from 2M NaCl to 500 mM NaCl, was performed at 4°C. The reconstitution apparatus and buffer were then pre-warmed to 37°C and at this stage GH1/GH5 was added. The remaining dialysis (500 mM to 80 mM) was carried out at 37 °C over a period of 3 hours. Chromatin prepared in this manner was then digested with MNase under standard conditions to produce populations of chromatosome and core particle DNAs (Figure 4-26A & 4-26B).

The results of this study of nucleosome positioning adopted during 37°C-reconstitution are presented in Figure 4-26C. In order to compare data, the same mapping results derived from a standard reconstitution (taken from Figure 4-17) are also presented and have been aligned to indicate corresponding boundaries. Figure

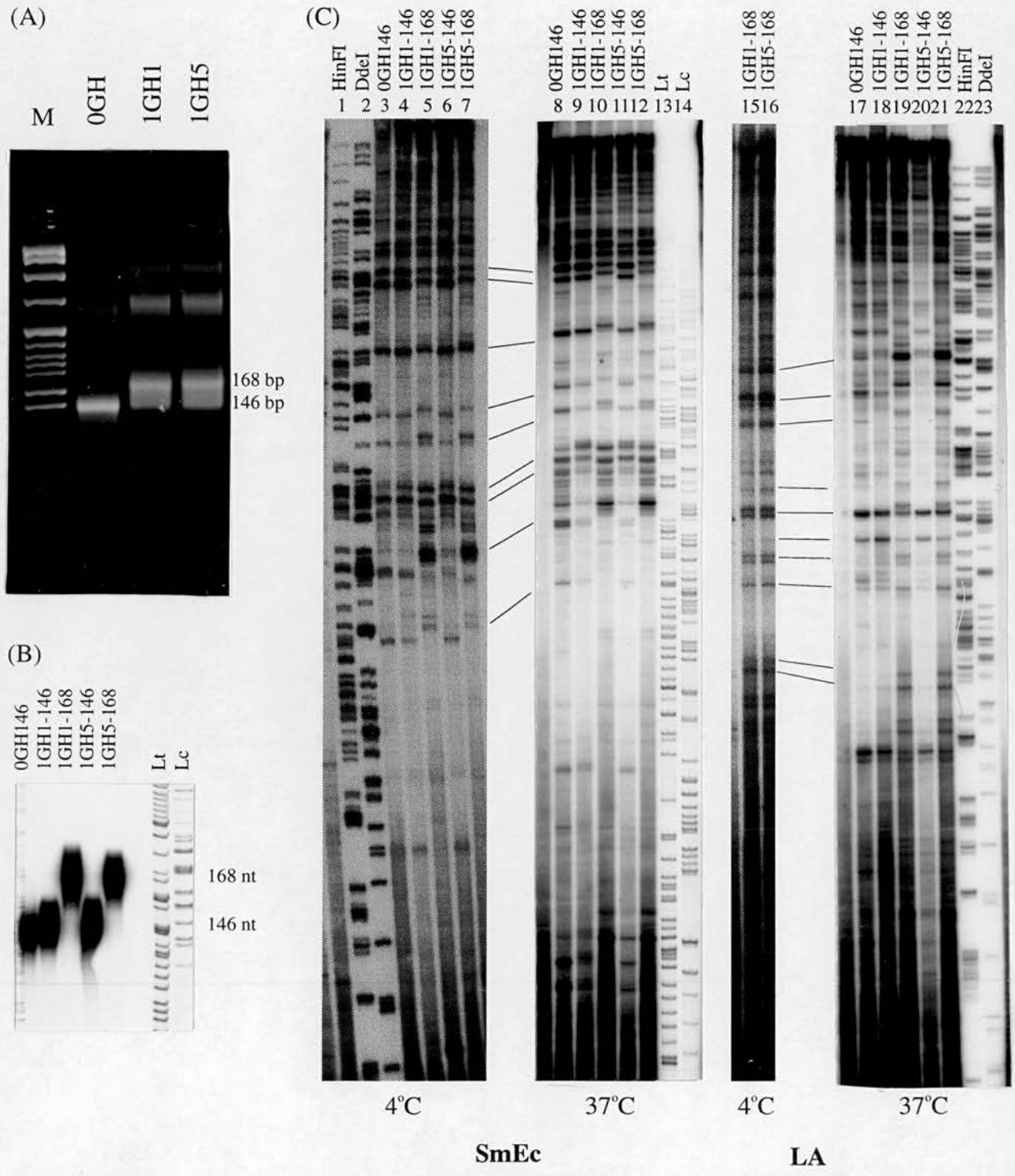


Figure 4-26 Comparison of nucleosome positions adopted during reconstitution at 4°C or at 37°C. (A) 4.5% metaphor agarose gel analysis of core and chromatosomal DNAs isolated from 37°C reconstitutes. (B) 6% denaturing polyacrylamide gel analysis of 5'-end labelled core particle and chromatosomal DNAs. (C) 6% denaturing polyacrylamide gel analyses of monomer extension products formed on mapping plasmids SmEc and LA. The data from 4°C reconstitutes (lanes 3 to 7 & 15, 16) are taken from the analysis shown in Figure 4-17. The data from 37 °C reconstitutes are shown in lanes 8 to 12 & 17 to 21. Corresponding bands derived from different reconstituted conditions have been indicated by lines.

4-26C shows that the core particle positioning derived from the 37°C reconstitution (lane 8) is very similar to the core particle positioning derived from the normal (4°C) reconstitution (lane 3). Also, the chromatosome positioning derived from 37°C-reconstitution (Figure 4-26C, lanes 10, 12 & 19, 21) is very similar to the chromatosome positioning derived from the 4°C reconstitution (Figure 4-26C, lanes 5, 7 & 15, 16).

The results also demonstrate that all the core particle positioning sites from the GH1- or GH5- containing chromatosome (Figure 4-26C, lanes 9, 18 & 11, 20) are very similar to the core particle positioning sites from the reconstitute which lacked linker histone globular domains (Figure 4-26C, lanes 8 & 17). Chromatosome positioning sites from GH1-containing chromatin (Figure 4-26, lanes 10 & 19) are very similar to chromatosome positioning site from GH5-containing chromatin (Figure 4-26, lanes 12 & 21).

These results suggest that core particle and chromatosome positioning is not influenced substantially by the temperature at which the final stages of reconstitution are carried out and by the salt concentration at which GH1/GH5 was added. This observation would appear to indicate that nucleosomes are either mobile or not mobile under both reconstitution processes.

4.3.6 Effects of the linker histones tails on the nucleosome positioning

Linker histones consist of a highly folded, conserved globular domain, a short, unstructured N-terminal tail and a long, basic, largely unstructured C-terminal tail. The different structural domains are thought to perform different roles in determining nucleosome and higher order chromatin structure. The globular domain seals off two turns of DNA in a nucleosome (Allan *et al.*, 1980), and the C-terminal tail interacts with linker DNA to permit neighbouring nucleosomes to come together and fold into higher order structure (Hill *et al.*, 1989). The function of the N-terminal tail of linker histones is still unclear, but it is believed that it may serve

some role in placing the globular domain precisely onto the nucleosome (Allan *et al.*, 1986).

It has been shown above that the addition of linker histone globular domains does not establish new nucleosome positions and that the chromatosome map is independent of the linker histone globular domain subtype. As the tail domains of different linker histone subtypes exhibit substantial variability compared to the conserved globular domain, it was relevant to investigate potential differences between (i) linker histone globular domains and intact linker histones and (ii) intact linker histone subtypes, with respect to their influence on nucleosome positioning. Therefore, monomer extension was again employed using reconstituted chromatin prepared with intact chicken erythrocyte H1 or H5. Core particle and chromatosome DNAs were prepared under standard MNase digestion conditions (Figure 4-27A).

As shown in Figure 4-27B, most of the chromatosome positioning sites derived from the H1 or H5-containing reconstitutes (lanes 5, 14 & 7, 16, respectively) are similar to the corresponding chromatosome positioning sites derived from the GH1 or GH5-containing reconstitutes (lanes 2, 11 & 3, 12). Although there are some minor quantitative differences between the corresponding digests, the patterns of positioning formed in the presence of globular domain or intact linker histone are very similar. Furthermore, the chromatosome positioning sites derived from H1-reconstitutes (Figure 4-27, lanes 5, 14) are very similar to the chromatosome positioning sites derived from H5-reconstitutes (Figure 4-27, lanes 7, 16).

These results suggest that the linker histone tails do not have a substantial influence upon nucleosome positioning in this system.

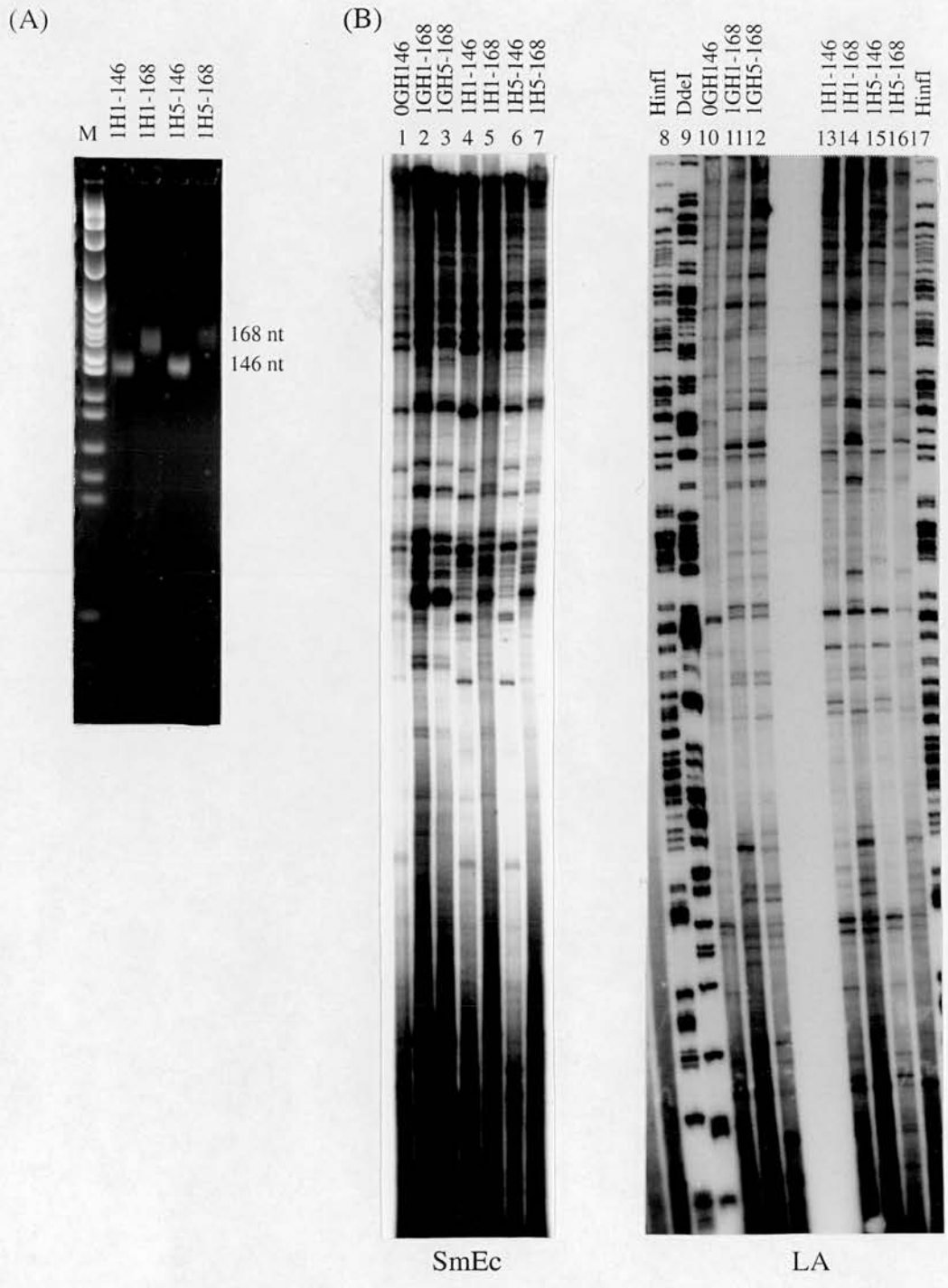


Figure 4-27 Comparison of effects of linker histone tails on nucleosome positioning. (A) 4.5% metaphor agarose gel analysis of core particle and chromatosomal DNAs from H1- or H5-containing reconstitutes. (B) 6% denaturing polyacrylamide gel analysis of monomer extension products formed on mapping constructs SmEc and LA. 1H1-146: core particle DNA derived from H1-containing reconstitutes. 1H1-168: chromatosomal DNA derived from H1-containing reconstitutes. 1H5-146: core particle DNA derived from H5-containing reconstitutes. 1H5-168: chromatosomal DNA derived from H5-containing reconstitutes. The other lane descriptions and size standards are as described in Figure 4-10 & 4-11. The marker (M) in (A) was an MspI digest of pBR322 DNA.

4.4 Discussion

Linker histones play a crucial role in inducing and maintaining higher order chromatin structure (Simpson, 1978; Allan *et al.*, 1980; Thoma *et al.*, 1983; Allan *et al.*, 1986; Hill *et al.*, 1989). The binding of the globular domain of linker histones to the nucleosome is a critical step in the pathway by which these molecules interact with chromatin and modulate its structure. The capacity of linker histone tails to fold and aggregate nucleosomes into higher order structure depends upon correct globular domain binding (Allan *et al.*, 1986). All nucleosomes are likely to interact with linker histones and clearly, therefore, this protein has major potential to influence the binding or the location of binding of core histones to the DNA. It has been suggested that higher order chromatin structure imposes a constraint on nucleosome positioning and therefore H1 can at least indirectly effect positioning as it is the agent of higher order folding. However, whether the binding of an H1 molecule to an individual nucleosome can alter its position is unclear. In this study, the effects of linker histones and their globular domains upon nucleosome positioning on a long stretch of chicken β-globin DNA (~1.5 kb) has been investigated *in vitro*. By employing the technique of monomer extension, I have gathered information which sheds light upon the manner in which linker histone globular domains interact with nucleosomes and how this capacity is conserved between subtypes and is influenced (or otherwise) by other linker histone domains.

4.4.1 Linker histones globular domains and nucleosome positioning

Many studies have revealed that DNA sequence is an important determinant of nucleosome positioning (Simpson, 1991; Thoma, 1992; Woodcock *et al.*, 1993). A nucleosome positioning map of the chicken β-globin gene has been determined. It identifies the major sites at which core histone octamers position on the DNA and provides information concerning the relative affinity of the octamer for each position (Davey *et al.*, 1995). The positioning of core histone octamers on a section of the

DNA studied by Davey *et al.* (1995) was re-examined in this study and found to be consistent (Figure 4-18 A & B), although there are minor quantitative differences between our observations which may be attributed to changes in mapping constructs or to variation in core particle preparation.

Although the core particle structure and the path of DNA on its surface have been determined at 2.8 Å resolution (Luger *et al.*, 1997), the location of the linker histone globular domain within a nucleosome is still uncertain. It is generally accepted that linker histones bind to the nucleosome and protect an extra 20 bp of linker DNA. This was first revealed by digestion of chromatin with micrococcal nuclease which identified the chromatosome, a nucleosome structure consisting of approximately 168 bp of DNA surrounding the histone octamer and one molecule of H1/H5 (Simpson, 1978). Allan *et al.* (1980) first showed that the linker histone globular domain, GH1 or GH5, interacts with DNA where it enters and exits the nucleosome after wrapping around the core histones, a position from where it could protect an additional 10 bp of DNA symmetrically distributed on either side of the dyad axis of the nucleosome core. Since then, it has been accepted that linker histone globular domains bind symmetrically to the nucleosome to protects an extra 10 bp of DNA on each end of the core particle.

By studying a nucleosome incorporating the somatic 5S rDNA gene from *Xenopus borealis*, Wolffe and his co-worker have suggested that GH5 is asymmetrically associated with the nucleosome core such that the extra protected DNA constitutes 15 bp on one side and 5 bp on the other side of the dyad (Hayes *et al.*, 1994). Based on a statistical analysis of 280 chromatosomal DNA sequences, Travers and Muyldermans (1996) also proposed an asymmetric location for linker histone globular domains on the nucleosome core. Binding was proposed to make contact with the extra DNA turn close to the dyad and another contact in the adjacent gyre close to one extremity. This arrangement could therefore generate 20 bp and 0 bp DNA extension in chromatosome protection.

Recently, Pruss *et al.* (1996) have probed the *Xenopus borealis* 5S rDNA positioned nucleosome, incorporating linker histone, with photo-activatable cross-linking reagents. They confirmed that GH5 is asymmetrically located and may be positioned inside the gyres of nucleosomal DNA. Subsequently, using a site-directed chemical mapping approach, Hayes (1996) also showed that the globular domain of $H1^0$ is asymmetrically located in the same 5S rDNA positioned nucleosome. A criticism of the above studies is that they all employed a single unique DNA fragment, the 5S rDNA gene sequence, for preparation of a positioned nucleosome such that it is possible that their observations are particular to that positioned nucleosome. In this study, it was possible to investigate the protection mode in a large population of chromatosomes. DNA extensions were determined for both upstream and downstream boundaries with respect to the corresponding core particle positioning sites over a 1.5 kb length of the β-globin gene. As shown in Table 4-1, 7 out of the 9 major core particle positioning sites displayed a symmetric mode of chromatosome protection. All of these had ~10 bp DNA extension at each end of the core structure. The exceptions were nucleosome 5B, which exhibited both 10+10 and 20+0 bp protection and nucleosome 7 which exhibited 15+5 bp protection.

Moreover, analysis of chromatosomal DNA extensions from a larger population of chromatosomes suggests that most of the chromatosomes adopt 10+10 bp extensions with respect to the protection from core particles. Although some other cases exhibited 15+5 bp and 20+0 bp asymmetrical protection, it can be calculated that the large majority of the population are expected to adopt 10+10 DNA extension (Figure 4-21A). The result has also been tested by application of the Null hypothesis and indicates that in a large population of chromatosomes most of the DNA extension should be 10+10 bp. These results strongly support the proposal that linker histone globular domains, in forming a chromatosome, generally protect DNA exiting from the core particle in a symmetrical manner.

The core particle positioning sites derived from GH1- or GH5-containing reconstitutes are very similar to the core particle positioning sites derived from the reconstitutes lacking linker histone globular domains, a feature which applies

throughout the mapping area and consistent through the various reconstitution modifications. Core particle positioning sites derived from GH1- or GH5-containing reconstitutes are produced by chromatosome degradation. If linker histone globular domains were establishing "new" or unique chromatosome positioning sites, they should not produce the equivalent core particle positioning site distribution. However, all the core particle positioning sites derived from reconstitutes containing GH1 or GH5 can be attributed to the core particle positioning sites derived from reconstitutes lacking linker histone globular domains. This suggests that the addition of linker histone globular domain does not change or establish positioning sites on the β-globin gene sequence under the reconstitution conditions employed here. Instead, chromatosomes are formed at sites determined by the specificity of core histone octamer binding. Furthermore, there is no difference between GH1 and GH5 in this context, and it appears that GH1 and GH5 have the same properties in response of influencing nucleosome positioning.

The crystal structure of GH5 has been solved to 2.5 Å (Ramakrishnan *et al.*, 1993) and the sequence of GH1 has been studied by NMR spectroscopy (Cerf *et al.*, 1994). Although subtle differences have been observed between the two structures and between the electrostatic potentials surrounding the molecules, both structures share the same three helices followed by a β-hairpin and are folded in a very similar manner. It is, therefore, reasonable to predict that GH1 and GH5 would display the same properties in relation to nucleosome binding and nucleosome positioning.

As the addition of linker histone globular domains does not alter positioning, strong chromatosome positioning sites should appear in accordance with corresponding strong core particle positioning sites. In most cases, extension products reflecting chromatosomes do display similar intensity to the corresponding core particle fragments. This indicates that chromatosome formation and stability is largely determined by core histone-DNA interactions. However, analysis in the nucleosome 2 cluster indicates that this is not always the case. Strong chromatosome bands can be derived from weak(er) core particle positioning sites. This indicates that the sequence arrangement of DNA at the positioning site, and particularly the

extra DNA involved in forming a chromatosome, may influence linker histone globular domain binding.

Although some chromatosomes appear to have 0 bp protected at one end, not all of these display the expected 20 bp protection at the other end. It is possible that the 0 bp protection found in these cases may result from internal nicking of chromatosomal DNA. The MNase digestion time course experiment showed that different chromatosomes can display different stabilities (Figure 4-23 & 4-24). Chromatosomes which accumulate during MNase digestion (positioning sites 2^{d}, 2^{e}, 2^{f} and 2^{h}) suggest a stable structure. In contrast, chromatosomes which are degraded very rapidly during digestion (positioning site 3) suggest a particularly unstable structure. Most nucleosomes fall between these two extremes, although one can still distinguish unstable (nucleosome 2^{m}) and average (2^{a}, 2^{b}, 2^{c}, 2^{g}, 2^{j}, 2^{k}, 2^{l}, 2^{n}, 4A and 4B) stability from the chromatosomes digestion.

It is possible that DNA sequence, or sequence-dependent chromatosome structure is responsible for variability in nucleosome sensitivity. If one considers the sequence at the downstream boundary of the nuclease-sensitive chromatosome site 3, it shows that the extra region protected by the linker histone globular domain is very AT-rich and includes an A_4 tract (Table 4-2). This AT-richness is also found for other unstable chromatosomes. For example, nucleosome positioning sites 2^{m}, 2^{n} and 4A, which displayed a moderately unstable behaviour, also contain AT-rich sequences in the region where the extra chromatosome DNA is joined to the core particle DNA (Table 4-2). These AT-rich sequences located within the end of the chromatosomal DNA may form some type of distortion which is detected by the nuclease. Alternatively, this sensitivity behaviour may simply reflect the known preference MNase has for AT-rich sequence. If one considers those chromatosomes with stable properties, it is notable that the extra, chromatosome-protected DNA is usually relatively GC-rich (Table 4-2, sites 2^{a} to 2^{l}).

		core	extension area
2[a]	5'-	CTGGC	AGCAGCCGT
2[b]	5'-	GCAGC	CGTGGCAGC
2[c]	5'-	TCCCA	GCACGCTGC
2[d]	5'-	CTGCC	ATGTCACCGGT
2[e]	5'-	CCATG	TCACCGGTCAGGTC
2[f]	5'-	GGTCG	GGTGCTGCCTTC
2[g]	5'-	CCTTC	CTGCT
2[h]	5'-	TGCTG	CCAGCCAGGCT
2[i]	5'-	CCAGC	CAGGCTGTCC
2[j]	5'-	CTGTC	CCTGGTGTACCC
2[k]	5'-	TGGTG	TACCCACTG
2[l]	5'-	CCACT	GTGTCCCACCT
2[m]	5'-	TCTGT	GTTCTGAG
2[n]	5'-	TGCAG	GATCTTT
3	5'-	TCAGC	AAAATGCTCA
4A	5'-	TGCTG	TGGTTTGGAA

Table 4-2 Extra sequence protected upon binding of the linker histone globular domain to selected nucleosomes.

4.4.2 The effects of linker histone globular domains on nucleosome positioning under condition optimised for nucleosome mobility

Pennings *et al.* (1991) have suggested a general property of mobility between nucleosome positioning sites, a feature which is temperature-sensitive. They argued that although nucleosome positioning can be precisely defined, core histones can migrate between these positioning sites (Meersseman *et al.*, 1991; Meersseman *et al.*, 1992). Moreover, this mobility is suppressed by the binding of H1 or H5 (Pennings *et al.*, 1994). The concept of mobile nucleosomes is attractive in that it has the potential to create transient accessibility for DNA-binding factors.

Under normal reconstitution conditions, gel analyses have shown that the core particle positioning sites derived from the GH1- or GH5-containing reconstitutes are very similar to the core particle positioning sites derived from the reconstitutes lacking the linker histone globular domains (Figures 4-11 to 4-17). These results suggest that the addition of linker histone globular domain onto preset core particles does not change their positioning. Furthermore, the core particle positioning sites derived from reconstitution conditions optimised to encourage nucleosome mobility are similar to the core particle positioning sites derived from the standard reconstitution conditions (Figure 4-26C). This demonstrates that nucleosome positioning is apparently indifferent to the reconstitution conditions and that the core histone octamer is the principal determinant of the position adopted. In principle, a higher temperature should provide a greater possibility for nucleosome mobility particularly when coupled to the introduction of GH1/GH5 at 500 mM Na^+. Whilst there is no formal demonstration of nucleosome mobility on this sequence, linker histones had no effect upon core histone octamer positioning under conditions at which nucleosome are predicted to be mobile (37°C), or at least, bound to the DNA but not as yet having adopted their final positions (37°C, high salt). The fact that no difference is observed in core and chromatosome positioning, either qualitative or quantitative, may suggest that nucleosomes are not mobile under both reconstitution conditions. In keeping with this latter possibility, Pennings

(unpublished) has found that histones octamers are noticeably lacking in mobility when reconstituted onto the 1.5 kb fragment analysed in this study.

4.4.3 Influence of linker histones on nucleosome positioning

It has been found that a polar, head-to-tail arrangement of H1/H5 molecules exists along extended chromatin (Ring & Cole, 1983; Lennard & Thomas, 1985). Moreover, when condensed, the neighbouring C-terminal tails are closer together. This arrangement suggests that the contact between C- and N-terminal tails has the potential to influence nucleosome positioning or structure by virtue of their interaction with neighbouring nucleosomes. Furthermore, the nucleosome filament has a higher affinity for H5 than H1 and this would lead to more stable H5-bound chromatin structure than H1-bound chromatin structure (Thomas & Rees, 1983). The reason for this difference in affinity may lie in their C-terminal tails, as H5 has a long, arginine-rich C-terminal tail whereas H1 has a slightly longer but lysine-rich C-terminal tail. Therefore, it is reasonable to expect that H1 might display different effects on nucleosome positioning from H5 because of these differences. Similarly, and perhaps more fundamentally, one might expect to see a difference between globular domains and intact molecules in this context.

However, this study has shown that the core particle positioning boundaries and chromatosome boundaries are very similar for H1 and H5 and indeed for GH1 and GH5 (Figure 4-27B). In the cell nucleus the repeat length of nucleosomes along the DNA varies from 165 bp to 250 bp. In chicken erythrocytes the average repeat is about 200 bp. Also, it has been demonstrated that in an *in vitro* reconstitution system, the addition of H5 or H1 can induce the formation of an ordered, spaced, arrangement of nucleosomes. However, this requires an appropriately high density of core histone octamers during reconstitution and can only be facilitated by the presence of polyglutamic acid (Stein & Kunzler, 1983; Stein & Bina, 1984). Both these factors were specifically avoided in this study. For monomer extension, chromatin was deliberately reconstituted at a core histone to DNA ratio of one core

histone octamer per 500 bp of DNA. Under this condition, the opportunity for linker histone-linker histone, or linker histone tail interactions may be very limited.

4.4.4 DNA sequence features and nucleosome positioning

Nucleosome positioning frequently depends on certain sequence patterns in the DNA. Intensive statistical data analysis has provided substantial evidence of sequence patterns involved in nucleosome positioning (Travers & Klug, 1987; Travers & Muyldermans, 1996; Ioshikhes *et al.*, 1996; Bolshoy *et al.*, 1996). Travers and his co-workers cloned and sequenced 280 DNA fragments isolated from chicken erythrocyte chromatosomes. They suggest that two DNA motifs, the GGA box and an AAA/TTT-rich region, could be involved in determining chromatosome positioning. They proposed that the short DNA sequence NGGR (where N=A, T, G, or C, and R= A or G) was usually located at one of the termini of the cloned DNA (Satchwell *et al.*, 1986; Muyldermans & Travers, 1994; Travers & Muyldermans, 1996). Using a multi-alphabet consensus algorithm, Ulyanov & Stormo (1995) have identified the patterns MMMNNMMM (where M=A or C) and RRRNNRRR as often being placed 40 bp and 15 bp away from the dyad axis, respectively. As nucleosome positioning sites display weak patterns, Trifonov and his colleagues have developed a multiple alignment algorithm to detect these weak patterns and found a ~10.4 bp oscillation occurrence of the dinucleotides AA and TT (Ioshikhes *et al.*, 1996; Bolshoy *et al.*, 1996). These results may indicate that certain DNA sequence motifs could present an important signal for positioning core histone octamers and chromatosomes.

In the light of such preferential occurrences of DNA motifs, the chromatosomes and core particles positioned on the 1.5 kb β-globin sequence were searched for these signals. The nine major chromatosome positioning sites were aligned at their dyad axes. The occurrences and position of eight different sequence motifs were studied. The midpoint of the nucleosome structure was denoted as 0 bp and extended 90 bp on each side. To simplify the analysis and to account for some

degree of inaccuracy or error, the appearance of the sequence motifs were grouped in 5 bp intervals. The results of this study are shown in Figure 4-28.

Generally speaking, there is little evidence from a visual inspection of the data to suggest that any of the motifs investigated play a substantial role in determining chromatosome positioning on the β-globin gene. However, a few of the analyses are worthy of some comment.

The motif RRRnnRRR (Figure 4-28F) shows a significant polar distribution, commonly occurring about 40 and 85 bp away from dyad at one side of the chromatosome only. This result is not compatible with the suggestion of Ulyanov & Stormo (1995) who placed the motif 15 bp away from the dyad axis.

The motif NGGR (Figure 4-28A) shows a detectable distribution at 5, 35, 55 and 85 bp away from the dyad. This result is inconsistent with the observation of Travers & Muyldermans (1996) who placed the motif at the very end of chromatosomal DNA.

The occurrences of trinucleotides AAA (Figure 4-28B) and TTT (Figure 4-28C) are relatively low and dinucleotides AA (Figure 4-28G) and TT (Figure 4-28H) do not show a 10.4 bp oscillation occurrence. Again, these results do not agree with previous predictions (Travers & Muyldermans, 1996; Ioshikhes *et al.*, 1996; Bolshoy *et al.*, 1996).

In summary, the positioning sites employed in this study, which are all strong positioning sites on the β-globin gene, do not seem to employ these particular DNA sequence motifs. Although mapped at high resolution, the limited number of chromatosomes analysed (9) limits statistical significant and may not have yielded more subtle positioning information.

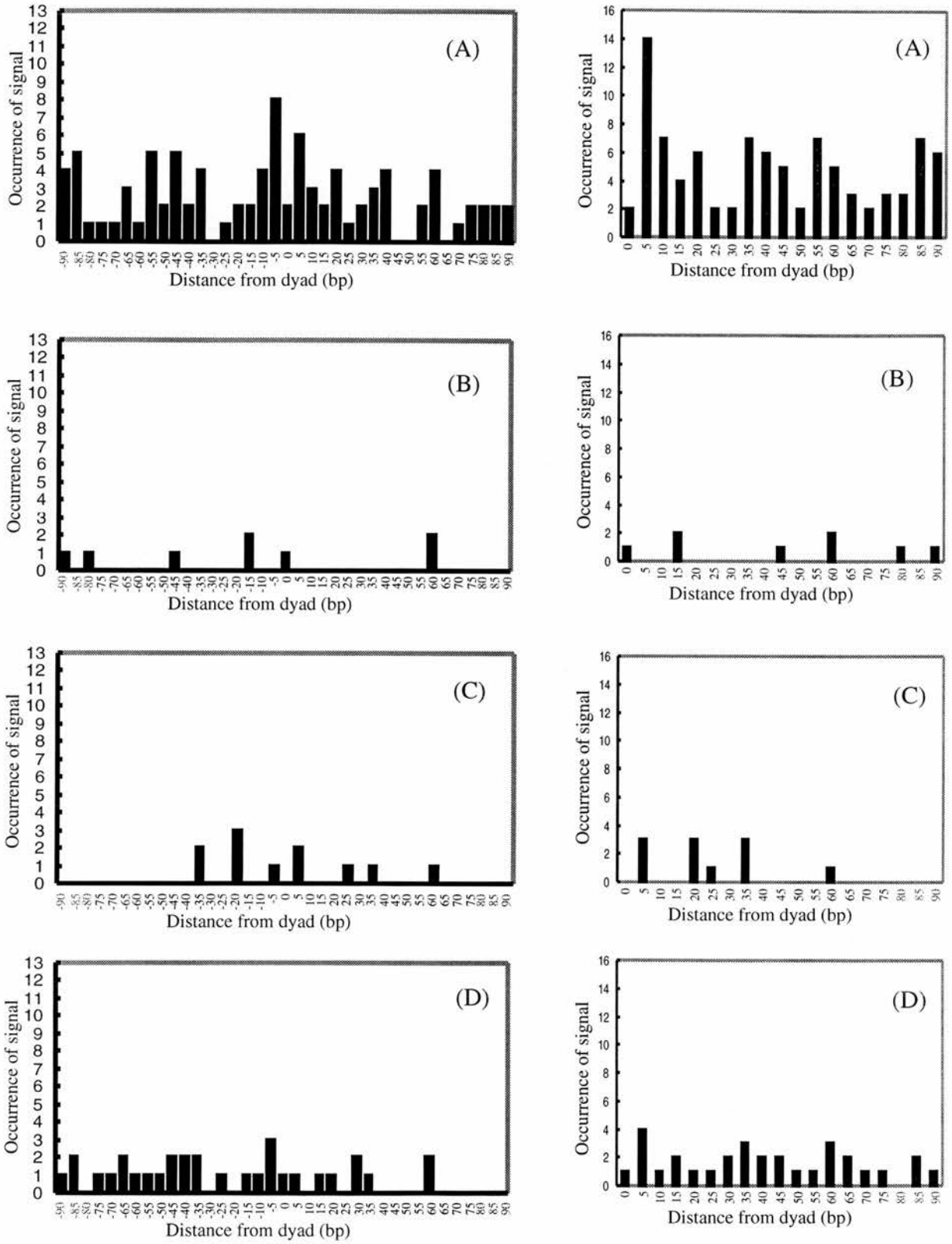
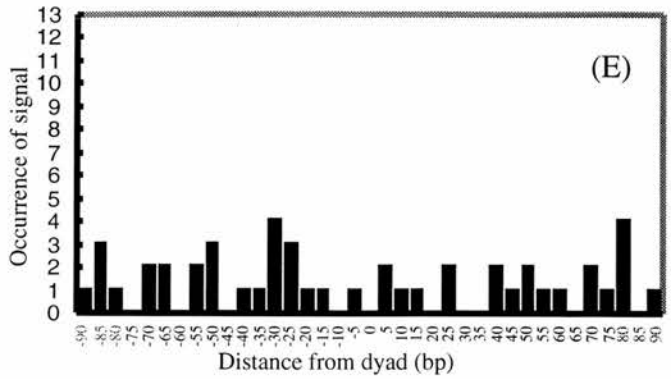
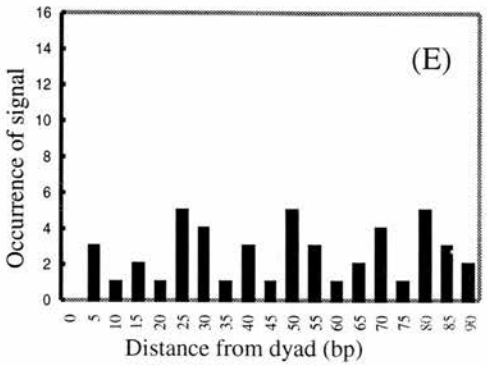
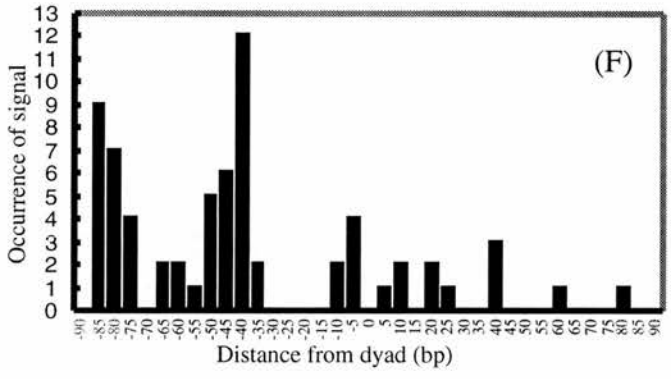
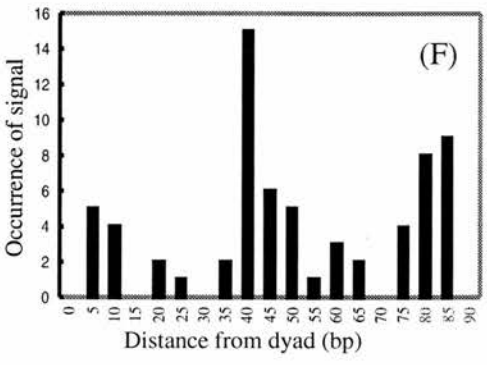


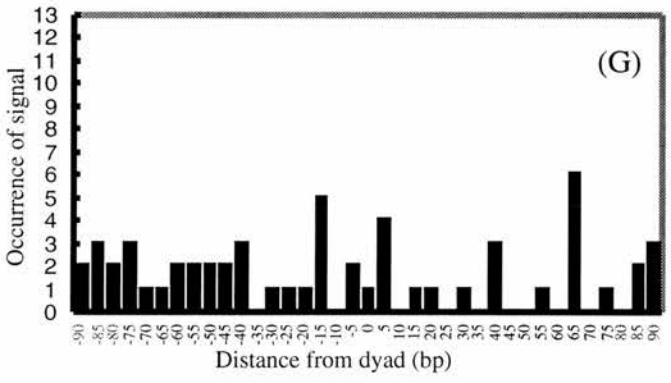
Figure 4-28 Occurrence of particular sequence motifs within chromatosome positioning sites. The sites used for this analysis were the 9 major positioning sites (1-3, 4A, 4B, 5A, 5B, 6 & 7). The histograms show the occurrence of the sequence motifs NGGR (A) , AAA (B) , TTT (C) , GGA (D) , MMMNNMMM (E), RRRNNRRR (F), AA (G) and TT (H) with respect to the centre (dyad axis) of the chromatosome DNA. The data have been grouped and are presented in 5 bp intervals. In the plots shown in the right hand panels, the data for the upstream and down stream occurrences have been superimposed so as to simply present the distance of the motifs from the dyad, irrespective of orientation.

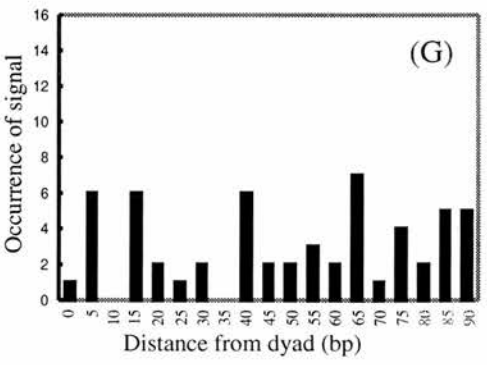


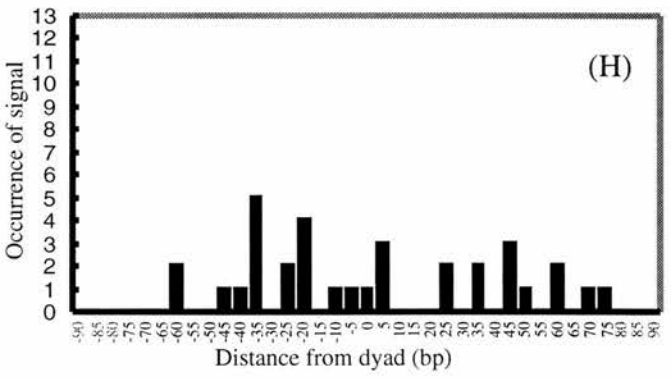












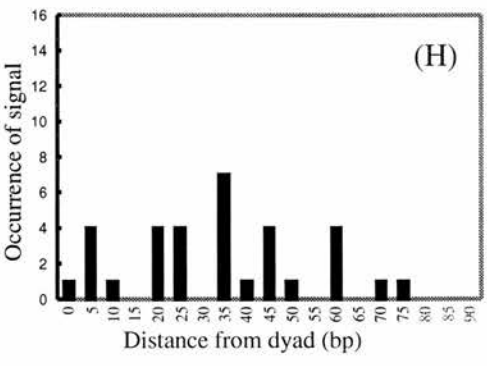


Figure 4-28 (cont.)

4.5 Concluding remarks

The central aims of this study were to assess the influence of linker histones and their globular domains on the formation and positioning of nucleosomes, and to elucidate any differences between subtypes in this context. Core particle positioning derived from reconstitutes containing GH1/GH5 or H1/H5 was found to be very similar to core particle positioning derived from reconstitutes lacking linker histones or their globular domains. Even under dynamic reconstitution conditions, the distribution of positioning sites was retained. These results strongly suggest that nucleosome positioning is not changed by the addition of linker histones or their globular domains. It appears, therefore, that nucleosome positioning is dominated by the interaction between the core histone octamer and its associated DNA sequence.

No significant difference was observed between linker histone subtypes in their affect on nucleosome positioning. Furthermore, there was little evidence to indicate a difference between intact linker histones and their globular domains in this respect. This suggests that the tails of linker histone do not contribute substantially to nucleosome positioning in the system employed here which may exclude the contribution of linker histone-linker histone interactions. To gain a full insight into the role of linker histones in nucleosome positioning and the determination of chromatin structure, further studies should aim to accommodate this feature of this important family of chromosomal proteins.

A major observation of this study concerns the distribution of the extra 20 bp of DNA that is associated with a chromatosome compared to a core particle. The results are consistent with a general theory for symmetric distribution (10+10 bp), but that exceptions to this do occur.

References

Adachi, Y., Luke, M. & Laemmli, U.K. (1991) Chromosome assembly *in vitro* – topoisomerase II is required for condensation. *Cell,* **64**, 137-148.

Adams, C.C. & Workman, J.L. (1993) Nucleosome displacement in transcription. *Cell,* **72**, 305-308.

Adams, C.C. & Workman, J.L. (1995) Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Mol Cell Biol,* **15**, 1405-1421.

Adolph, K.W., Cheng, S.M. & Laemmli, U.K. (1977) Role of nonhistone proteins in metaphase chromosome structure. *Cell,* **12**, 805

Aizawa, S., Nishino, H., Saito, K., Kimura, K., Shirakawa, H. & Yoshida, M. (1994) Stimulation of transcription in cultured-cells by high-mobility group protein-1 - essential role of the acidic carboxyl-terminal region. *Biochemistry,* **33**, 14690-14695.

Albright, S.C., Wiseman, J.M., Lange, R.A. & Garrard, W.T. (1980) Subunit structures of different electrophoretic forms of nucleosomes. *J Biol Chem,* **255**, 3673-3684.

Alfonso, P.J., Crippa, M.P., Hayes, J.J. & Bustin, M. (1994) The footprint of chromosomal-proteins HMG-14 and HMG-17 on chromatin subunits. *J Mol Biol,* **236**, 189-198.

Allan, J., Hartman, P.G., Crane-Robinson, C. & Aviles, F.X. (1980) The structure of histone H1 and its location in chromatin. *Nature,* **288**, 675-679.

Allan, J., Mitchell, T., Harborne, N., Bohm, L. & Crane-Robinson, C. (1986) Roles of H1 domains in determining higher-order chromatin structure and H1 location. *J Mol Biol,* **187**, 591-601.

Almer, A., Rudolph, H., Hinnen, A. & Horz, W. (1986) Removal of positioned nucleosomes from the yeast pho5 promoter upon pho5 induction releases additional upstream activating DNA elements. *EMBO J,* **5**, 2689-2696.

Archer, T.K., Cordingley, M.G., Wolford, R.G. & Hager, G.L. (1991) Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary-tumor virus promoter. *Mol Cell Biol,* **11**, 688-698.

Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E. & Moudrianakis, E.N. (1991) The nucleosomal core histone octamer at 3.1Å resolution - a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci USA,* **88**, 10148-10152.

Arents, G. & Moudrianakis, E.N. (1993) Topography of the histone octamer surface - repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc Natl Acad Sci USA,* **90**, 10489-10493.

Arents, G. & Moudrianakis, E.N. (1995) The histone fold - a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc Natl Acad Sci USA,* **92**, 11170-11174.

Aubert, D., Garcia, M., Benchaibi, M., Poncet, D., Chebloune, Y., Verdier, G., Nigon, V., Samarut, J. & Mura, C.V. (1991) Inhibition of proliferation of primary avian fibroblasts through expression of histone H5 depends on the degree of phosphorylation of the protein. *J Cell Biol,* **113**, 497-506.

Aubin, R.J., Frechette, A., De-Murcia, G., Mandel, P., Lord, A., Grondin, G. & Poirier, G.G. (1983) Correlation between endogenous nucleosomal hyper(ADP-ribosyl)ation of histone H1 and the induction of chromatin relaxation. *EMBO J,* **2**, 1685-1693.

Ausio, J., Dong, F. & van Holde, K.E. (1989) Use of selectively trypsinized nucleosome core particles to analyze the role of the histone tails in the stabilization of the nucleosome. *J Mol Biol,* **206**, 451-463.

Aviles, J., Chapman, G.E., Kneale, G.G., Crane-Robinson, C. & Bradbury, E.M. (1978) The conformation of histone H5. *Eur J Biochem,* **88**, 363-371.

Baldi, P., Brunak, S., Chauvin, Y. & Krogh, A. (1996) Naturally occurring nucleosome positioning signals in human exons and introns. *J Mol Biol,* **263**, 503-510.

Bannister, A.J. & Kouzarides, T. (1996) The cbp coactivator is a histone acetyltransferase. *Nature,* **384**, 641-643.

Barlev, N.A., Candau, R., Wang, L.A., Darpino, P., Silverman, N. & Berger, S.L. (1995) Characterization of physical interactions of the putative transcriptional adapter, Ada2, with acidic activation domains and TATA-binding protein. *J Biol Chem,* **270**, 19337-19344.

Barrick, D., Villanueba, K., Childs, J., Kalil, R., Schneider, T.D., Lawrence, C.E., Gold, L. & Stormo, G.D. (1994) Quantitative analysis of ribosome binding sites in *Escherichia coli. Nucl Acids Res,* **22**, 1287-1295.

Baumgartner, M., Dutrillaux, B., Lemieux, N., Lillenbaum, A. & Paulin, D. (1991) Gene occupy a fixed and symmetrical position on sister chromatids. *Cell,* **64**, 761-766.

Bavykin, S.G., Usachenko, S.I., Zalensky, A.O. & Mirzabekov, A.D. (1990) Structure of nucleosomes and organization of internucleosomal DNA in chromatin. *J Mol Biol,* **212**, 495-511.

Beard, P. (1978) Mobility of histones on the chromatin of simian virus 40. *Cell,* **15**, 955-967.

Becker, P.B. & Wu, C. (1992) Cell-free system for assembly of transcriptionally repressed chromatin from *Drosophila* embryos. *Mol Cell Biol,* **12**, 2241-2249.

Benton, B.M., Eng, W.K., Dunn, J.J., Studier, F.W., Sternglanz, R. & Fisher, P.A. (1990) Signal-mediated import of bacteriophage T7 RNA polymerase into the *Saccharomyces cerevisiae* nucleus and specific transcription of target genes. *Mol Cell Biol,* **10**, 353-360.

Bergman, L.W. (1986) A 278 bp DNA fragment containing the upstream activator sequence determines precise nucleosome positioning of the yeast pho5 gene. *J Cell Biochem,* 111

Bernardi, F., Zatchej, M. & Thoma, F. (1992) Species-specific protein - DNA interactions may determine the chromatin units of genes in *Saccharomyces cerevisiae* and in s-pombe. *EMBO J,* **11**, 1177-1185.

Bina, M. (1994) Periodicity of dinucleotides in nucleosomes derived from simian virus 40 chromatin. *J Mol Biol,* **235**, 198-208.

Bingham, A.H.A. & Busby, S.J.W. (1987) Translation of gale and coordination of galactose operon expression in *Escherichia coli* - effects of insertions and deletions in the non- translated leader sequence. *Mol Microbiol,* **1**, 117-124.

Bock, H., Abler, S., Zhang, X.Y., Fritton, H. & Igo-Kemenes, T. (1984) Positioning of nucleosomes in satellite I-containing chromatin of rat liver. *J Mol Biol,* **176**, 131-154.

Bolshoy, A., Mcnamara, P., Harrington, R.E. & Trifonov, E.N. (1991) Curved DNA without AA - experimental estimation of all 16 DNA wedge angles. *Proc Natl Acad Sci USA,* **88**, 2312-2316.

Bolshoy, A. (1995) CC dinucleotides contribute to the bending of DNA in chromatin. *Nature Struc Biol,* **2**, 446-448.

Bolshoy, A., Ioshikhes, I. & Trifonov, E.N. (1996) Applicability of the multiple alignment algorithm for detection of weak patterns: periodically distributed DNA pattern as a study case. *CABIOS,* **12**, 383-389.

Borodovsky, M. & Peresetsky, A. (1994) Deriving non-homogenous DNA Markov chain models by cluster analysis algorithm minimizing multiple alignment entropy. *Comput Chem,* **18**, 259-267.

Bouvet, P., Dimitrov, S. & Wolffe, A.P. (1994) Specific regulation of *Xenopus* chromosomal 5S rRNA gene transcription *in vivo* by histone H1. *Genes Dev,* **8**, 1147-1159.

Boy de la Tour, E. & Laemmli, U.K. (1988) The metaphase scaffold is helically folded: sister chromatids have predominantly opposite helical handedness. *Cell,* **55**, 937-944.

Bradbury, E.M. (1992) Reversible histone modifications and the chromosome cell-cycle. *BioEssays,* **14**, 9-16.

Bradbury, E.M., Inglis, R.J. & Matthews, H.R. (1974) Control of cell division by very lysine rich histone phosphorylation. *Nature,* **247**, 257-261.

Brandt, W.F. & VonHolt, C. (1986) Variants of wheat histone H1 with N-terminal and C-terminal extensions. *FEBS Lett,* **194**, 282-286.

Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D. & Broach, J.R. (1993) Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev,* **7**, 592-604.

Bresnick, E.H., Bustin, M., Marsaud, V., Richardfoy, H. & Hager, G.L. (1992) The transcriptionally active MMTV promoter is depleted of histone H1. *Nucl Acids Res,* **20**, 273-278.

Brownell, J.E. & Allis, C.D. (1996) Special hats for special occasions - linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev,* **6**, 176-184.

Brukner, I., Sanchez, R., Suck, D. & Pongor, S. (1995a) Trinucleotide models for DNA bending propensity - comparison of models based on DNaseI digestion and nucleosome packaging data. *J Biomol Struc Dynam,* **13**, 309-317.

Brukner, I., Sanchez, R., Suck, D. & Pongor, S. (1995b) Sequence-dependent bending propensity of DNA as revealed by DNaseI parameters for trinucleotides. *EMBO J,* **14**, 1812-1818.

Buckle, R., Balmer, M., Yenidunya, A. & Allan, J. (1991) The promoter and enhancer of the inactive chicken β-globin gene contains precisely positioned nucleosomes. *Nucl Acids Res,* **19**, 1219-1226.

Buckle, R.S., Maman, J.D. & Allan, J. (1992) Site-directed mutagenesis studies on the binding of the globular domain of linker histone H5 to the nucleosome. *J Mol Biol,* **223**, 651-659.

Buell, G., Schulz, M.F., Selzer, G., Chollet, A., Movva, N.R., Semon, D., Escanez, S. & Kawashima, E. (1985) Optimizing the expression in *E.coli* of a synthetic gene encoding somatomedin-C (IGF-1). *Nucleic Acids Res,* **13**, 1923-1938.

Butler, P.J.G. (1984) A defined structure of the 30-nm chromatin fiber which accommodates different nucleosomal repeat lengths. *EMBO J,* **3**, 2599-2604.

Cacchione, S., De Santis, P., Foti, D., Palleschi, A. & Savino, M. (1989) Periodical polydeoxynucleotides and DNA curvature. *Biochemistry,* **28**, 8706-8713.

Cacchione, S., Cerone, M.A., De Santis, P. & Savino, M. (1995) Superstructural features of the upstream regulatory regions of 2 pea rbcs genes and nucleosomes positioning - theoretical prediction and experimental evaluation. *Biophys Chem,* **53**, 267-281.

Calladine, C.R. & Drew, H.R. (1986) Principles of sequence-dependent flexure of DNA. *J Mol Biol,* **192**, 907-918.

Calladine, C.R., Drew, H.R. & McCall, M.J. (1988) The intrinsic curvature of DNA in solution. *J Mol Biol,* **201**, 127-137.

Calladine, C.R. & Drew, H.R. (1996) A useful role for static models in elucidating the behavior of DNA in solution. *J Mol Biol,* **257**, 479-485.

Callan, H.G. (1986) Lampbrush chromosomes. *Springer-Verlag*, Berlin,

Calogero, R.A., Pon, C.L., Canonaco, M.A. & Gualerzi, C.O. (1988) Selection of the mRNA translation initiation region by *Escherichia coli* ribosomes. *Proc Natl Acad Sci USA,* **85**, 6427-6431.

Canonaco, M.A., Gualerzi, C.O. & Pon, C.L. (1989) Alternative occupancy of a dual ribosomal-binding site by mRNA affected by translation initiation factors. *Eur J Biochem,* **182**, 501-506.

Caplan, A., Kimura, T., Gould, H. & Allan, J. (1987) Perturbation of chromatin structure in the region of the adult β- globin gene in chicken erythrocyte chromatin. *J Mol Biol,* **193**, 57-69.

Caron, F. & Thomas, J.O. (1981) Exchange of histone H1 between segments of chromatin. *J Mol Biol,* **146**, 513-537.

Cartwright, I.L., Keene, M.A. & Elgin, S.C.R. (1983) Site-specific chromatin patterns in *Drosophila melanogaster. J Cell Biol,* **97**, A389

Cattini, P.A., Harborne, N. & Allan, J. (1988) The linker histone stoichiometry of chicken erythrocyte chromatin *in situ. J Electron Microscopy,* **37**, 31-37.

Cerf, C., Lippens, G., Ramakrishnan, V., Muyldermans, S., Segers, A., Wyns, L., Wodak, S.J. & Hallenga, K. (1994) Homonuclear and heteronuclear 2-dimensional NMR studies of the globular domain of histone H1 - full assignment, tertiary

structure, and comparison with the globular domain of histone H5. *Biochemistry,* **33**, 11079-11086.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. (1994) Green fluorescent protein as a marker for gene expression. *Science,* **263**, 802-805.

Chasman, D.I., Lue, N.F., Buchman, A.R., Lapointe, J.W., Lorch, Y. & Kornberg, R.D. (1990) A yeast protein that influences the chromatin structure of uasg and functions as a powerful auxiliary gene activator. *Genes Dev,* **4**, 503-514.

Chen, C.Y. & Sarnow, P. (1995) Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science,* **268**, 415-417.

Chen, H., Li, B.Y. & Workman, J.L. (1994) A histone-binding protein, nucleoplasmin, stimulates transcription factor-binding to nucleosomes and factor-induced nucleosome disassembly. *EMBO J,* **13**, 380-390.

Chen, H.Y., Bjerknes, M., Kumar, R. & Jay, E. (1994) Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucl Acids Res,* **22**, 4953-4957.

Choi, O.R. & Engel, J.D. (1986) A 3′ enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult β-globin gene. *Nature,* **323**, 731-734.

Chuang, P.T., Albertson, D.G. & Meyer, B.J. (1994) DPY-27: a chromosome condensation protein homolog that regulate *C. elegans* dosage compensation through association with the X chromosome. *Cell,* **79**, 459-474.

Churchill, M.E.A. & Suzuki, M. (1989) 'SPKK' motifs prefer to bind to DNA at A/T-rich sites. *EMBO J,* **8**, 4189-4195.

Churchill, M.E.A. & Travers, A.A. (1991) Protein motifs that recognize structural features of DNA. *Trends Biochem Sci,* **16**, 92-97.

Clore, G.M., Gronenborn, A.M., Nilges, M., Sukumaran, D.K. & Zarbock, J. (1987) The polypeptide fold of the globular domain of histone H5 in solution - a study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO J,* **6**, 1833-1842.

Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G. & Ward, W.W. (1993) Chemical structure of the hexapeptide chromophore of the *Aequorea* green fluorescent protein. *Biochemistry,* **32**, 1212-1218.

Cole, R.D. (1984) A minireview of microheterogeneity in H1 histone and its possible significance. *Anal Biochem,* **136**, 24-30.

Cone, K.C. & Steege, D.A. (1985a) Functional analysis of lac repressor restart sites in translational initiation and reinitiation. *J Mol Biol,* **186**, 733-742.

Cone, K.C. & Steege, D.A. (1985b) mRNA conformation and ribosome selection of translational reinitiation sites in the lac repressor mRNA. *J Mol Biol,* **186**, 725-732.

Cormack, B.P., Valdivia, R.H. & Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene,* **173**, 33-38.

Costanzo, G., Dimauro, E., Salina, G. & Negri, R. (1990) Attraction, phasing and neighbor effects of histone octamers on curved DNA. *J Mol Biol,* **216**, 363-374.

Cote, J., Quinn, J., Workman, J.L. & Peterson, C.L. (1994) Stimulation of Gal4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science,* **265**, 53-60.

Crane-Robinson, C. & Ptitsyn, O.B. (1989) Binding of the globular domain of linker histones H5/H1 to the nucleosome - a hypothesis. *Protein Engineering,* **2**, 577-582.

Crane-Robinson, C. (1997) Where is the globular domain of linker histone located on the nucleosome? *Trends Biochem Sci,* **22**, 75-77.

Crothers, D.M., Haran, T.E. & Nadeau, J.G. (1990) Intrinsically bent DNA. *J Biol Chem,* **265**, 7093-7096.

Csordas, A. (1990) On the biological role of histone acetylation. *Biochem J,* **265**, 23-38.

Dasso, M., Dimitrov, S. & Wolffe, A.P. (1994) Nuclear assembly is independent of linker histones. *Proc Natl Acad Sci USA,* **91**, 12477-12481.

Davey, C., Pennings, S., Meersseman, G., Wess, T.J. & Allan, J. (1995) Periodicity of strong nucleosome positioning sites around the chicken adult β-globin gene may encode regularly spaced chromatin. *Proc Natl Acad Sci USA,* **92**, 11210-11214.

Davey, C., Pennings, S. & Allan, J. (1997) CpG methylation remodels chromatin structure *in vitro*. *J Mol Biol,* **267**, 276-288.

Davie, J.R. (1996) Histone modifications, chromatin structure, and the nuclear matrix. *J Cell Biochem,* **62**, 149-157.

DeLange, R.J., Fambrough, D.M., Smith, E.L. & Bonner, J. (1969) Calf and pea histone IV. III. complete amino acid sequence of pea seedling histone IV: comparison with the homologous calf thymus histone. *J Biol Chem,* **244**, 5669-5679.

De-Murcia, G., Huletsky, A., Lamarre, D., Gaudreau, A., Pouyet, J., Daune, M. & Poirier, G.G. (1986) Modulation of chromatin superstructure induced by poly(ADP-ribose) synthesis and degradation. *J Biol Chem,* **261**, 7011-7017.

De Santis, P., Palleschi, A., Savino, M. & Scipioni, A. (1990) Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry,* **29**, 9269-9273.

De Santis, P., Palleschi, A., Savino, M. & Scipioni, A. (1992) Theoretical prediction of the gel electrophoretic retardation changes due to point mutations in a tract of sv40 DNA. *Biophys Chem,* **42**, 147-152.

De Santis, P., Fua, M., Palleschi, A. & Savino, M. (1993) Relationships between intrinsic and induced curvature in DNAs -theoretical prediction of nucleosome positioning. *Biophys Chem,* **46**, 193-204.

De Santis, P., Fua, M., Palleschi, A. & Savino, M. (1995) Influence of dynamic fluctuations on DNA curvature. *Biophys Chem,* **55**, 261-271.

De Smit, M.H. & Van Duin, J. (1990a) Control of prokaryotic translational initiation by mRNA secondary structure. *Prog Nucl Acid Res Mol Biol,* **38**, 1-35.

De Smit, M.H. & Van Duin, J. (1990b) Secondary structure of the ribosome binding site determines translational efficiency - a quantitative-analysis. *Proc Natl Acad Sci USA,* **87**, 7668-7672.

De Smit, M.H. & Van Duin, J. (1994b) Control of translation by mRNA secondary structure in *Escherichia coli* - a quantitative-analysis of literature data. *J Mol Biol,* **244**, 144-150.

De Smit, M.H. & Van Duin, J. (1994a) Translational initiation on structured messengers - another role for the Shine-Dalgarno interaction. *J Mol Biol,* **235**, 173-184.

Diekmann, S. (1987) Temperature and salt dependence of the gel migration anomaly of curved DNA fragments. *Nucl Acids Res,* **15**, 247-265.

Diekmann, S. & Mclaughlin, L.W. (1988) DNA curvature in native and modified EcoRI recognition sites and possible influence upon the endonuclease cleavage reaction. *J Mol Biol,* **202**, 823-834.

Dimitrov, S.I., Russanova, V.R. & Pashev, I.G. (1987) The globular domain of histone H5 is internally located in the 30-nm chromatin fiber - an immunochemical study. *EMBO J,* **6**, 2387-2392.

Dimitrov, S. & Wolffe, A.P. (1996) Remodeling somatic nuclei in *Xenopus laevis* egg extracts -molecular mechanisms for the selective release of histones H1 and H1-degrees from chromatin and the acquisition of transcriptional competence. *EMBO J,* **15**, 5897-5906.

Dlakic, M. & Harrington, R.E. (1996) The effects of sequence context on DNA curvature. *Proc Natl Acad Sci USA,* **93**, 3847-3852.

Drew, H.R. & Travers, A.A. (1985) DNA bending and its relation to nucleosome positioning. *J Mol Biol,* **186**, 773-790.

Drew, H.R. & Calladine, C.R. (1987) Sequence-specific positioning of core histones on an 860 base-pair DNA - experiment and theory. *J Mol Biol,* **195**, 143-173.

Dreyfus, M. (1988) What constitutes the signal for the initiation of protein synthesis on *Escherichia coli* mRNAs? *J Mol Biol,* **204**, 79-94.

Du, W., Thanos, D. & Maniatis, T. (1993) Mechanisms of transcriptional synergism between distinct virus-inducible enhancer elements. *Cell,* **74**, 887-898.

Dunphy, W.G. & Newport, J.W. (1988) Mitosis inducing factors are present in a latent form during interphase in the *Xenopus* embryo. *J Cell Biol,* **106**, 2047-2056.

Durrin, L.K., Mann, R.K., Kayne, P.S. & Grunstein, M. (1991) Yeast histone H4 N-terminal sequence is required for promoter activation *in vivo*. *Cell,* **65**, 1023-1031.

Edmondson, D.G., Smith, M.M. & Roth, S.Y. (1996) Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev,* **10**, 1247-1259.

Eickbusch, T.H. & Moudrianakis, E.N. (1978) Compaction of DNA helices into either continuous supercoils or folded-fiber rods and toroids. *Cell,* **13**, 295-306.

Eisfeld, K., Candau, R., Truss, M. & Beato, M. (1997) Binding of NF1 to the MMTVpromoter in nucleosomes: influence of rotational phasing, translational positioning and histone H1. *Nucl Acids Res,* **25**, 3733-3742.

Elfring, L.K., Deuring, R., McCallum, C.M., Peterson, C.L. & Tamkun, J.W. (1994) Identification and characterization of *Drosophila* relatives of the yeast transcriptional activator SNF2/SWI2. *Mol Cell Biol,* **14**, 2225-2234.

Elgin, S.C.R. (1988) The formation and function of DNaseI hypersensitive sites in the process of gene activation. *J Biol Chem,* **263**, 19259-19262.

Falk, G. & Walker, J.E. (1988) DNA sequence of a gene cluster coding for subunits of the f0 membrane sector of ATP synthase in rhodospirillum-rubrum -support for modular evolution of the f1 and f0 sectors. *Biochem J.* **254**, 109-122.

Faxen, M., Plumbridge, J. & Isaksson, L.A. (1991) Codon choice and potential complementarity between mRNA downstream of the initiation codon and bases 1471-1480 in 16S rRNA affects expression of glns. *Nucl Acids Res,* **19**, 5247-5251.

Fedor, M.J., Lue, N.F. & Kornberg, R.D. (1988) Statistical positioning of nucleosomes by specific protein binding to an upstream activating sequence in yeast. *J Mol Biol,* **204**, 109-127.

Felsenfeld, G. (1992) Chromatin as an essential part of the transcriptional mechanism. *Nature,* **355**, 219-224.

Finch, J.T., Noll, M. & Kornberg, R.D. (1975) Electron microscopy of defined length of chromatin. *Proc Natl Acad Sci USA,* **72**, 3320-3322.

Finch, J.T. & Klug, A. (1976) Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci USA,* **73**, 1897-1901.

Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M. & Klug, A. (1977) Structure of nucleosome core particles of chromatin. *Nature,* **269**, 29-36.

Flaus, A., Luger, K., Tan, S. & Richmond, T.J. (1996) Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc Natl Acad Sci USA,* **93**, 1370-1375.

Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. & Turner, D.H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA,* **83**, 9373-9377.

Ganoza, M.C., Kofoid, E.C., Marliere, P. & Louis, B.G. (1987) Potential secondary structure at translation initiation sites. *Nucl Acids Res,* **15**, 345-360.

Garciaramirez, M., Rocchini, C. & Ausio, J. (1995) Modulation of chromatin folding by histone acetylation. *J Biol Chem,* **270**, 17923-17928.

Gasser, S.M. & Laemmli, U.K. (1986) The organization of chromatin loops characterization of a scaffold attachment site. *EMBO J,* **5**, 511-518.

Georgakopoulos, T., Gounalaki, N. & Thireos, G. (1995) Genetic-evidence for the interaction of the yeast transcriptional coactivator proteins Gcn5 and Ada2. *Mol Gen Genet,* **246**, 723-728.

Georgiou, G. (1996) Expression of proteins in bacteria. In: *Protein engineering: principles and practice*, edited by Cleland, J.L. & Craik, C.S. Wiley Liss, New York, p. 101-127.

Gerchman, S.E., Graziano, V. & Ramakrishnan, V. (1994) Expression of chicken linker histones in *Escherichia coli* -sources of problems and methods for overcoming some of the difficulties. *Protein Expression Purif,* **5**, 242-251.

Gheysen, D., Iserentant, D., Derom, C. & Fiers, W. (1982) Systematic alteration of the nucleotide sequence preceding the translation initiation codon and the effects on bacterial expression of the cloned SV40 small T antigen gene. *Gene,* **17**, 55-63.

Godde, J.S., Nakatani, Y. & Wolffe, A.P. (1995) The amino terminal tails of the core histones and the translational position of the TATA box determine TBP/TFIIA association with nucleosomal DNA. *Nucl Acids Res,* **23**, 4557-4564.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S. & Stormo, G. (1981) Translational initiation in prokaryotes. *Annu Rev Microbiol,* **35**, 365-403.

Gold, L., Stormo, G.D. & Saunders, R. (1984) *Escherichia coli* translational initiation factor IF3: a unique case of translational regulation. *Proc Natl Acad Sci USA,* **81**, 7061-7065.

Gold, L. (1988) Posttranscriptional regulatory mechanisms in *Escherichia coli. Ann Rev Biochem,* **57**, 199-233.

Gold, L. (1990) Expression of heterologous proteins in *Escherichia coli. Methods Enzymol,* **185**, 11-14.

Goodsell, D.S. & Dickerson, R.E. (1994) Bending and curvature calculations in B-DNA. *Nucl Acids Res,* **22**, 5497-5503.

Goytisolo, F.A., Gerchman, S.E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V. & Thomas, J.O. (1996) Identification of 2 DNA-binding sites on the globular domain of histone H5. *EMBO J,* **15**, 3421-3429.

Graifer, D.M., Malygin, A.A., Matasova, N.B., Mundus, D.A., Zenkova, M.A. & Karpova, G.G. (1997) Studying functional significance of the sequence 980-1061 in the central domain of human 18S rRNA using complementary DNA probes. *Biochim Biophys Acta-Gene Structure And Expression,* **1350**, 335-344.

Graziano, V., Gerchman, S.E., Schneider, D.K. & Ramakrishnan, V. (1994) Histone H1 is located in the interior of the chromatin 30-nm filament. *Nature,* **368**, 351-354.

Green, G.R., Lee, H.J. & Poccia, D.L. (1993) Phosphorylation weakens DNA-binding by peptides containing multiple SPKK sequences. *J Biol Chem,* **268**, 11247-11255.

Gren, E.J. (1984) Recognition of messenger RNA during translational initiation in *Escherichia coli. Biochimie,* **66**, 1-29.

Griffith, J., Bleyman, M., Rauch, C.A., Kitchin, P.A. & Englund, P.T. (1986) Visualization of the bent helix in kinetoplast DNA by electron microscopy. *Cell,* **46**, 717-724.

Groeneveld, H., Thimon, K. & Van Duin, J. (1995) Translational control of maturation protein synthesis in phage MS2 - a role for the kinetics of RNA folding. *RNA-A Publication Of The RNA Society,* **1**, 79-88.

Grosjean, H. & Fiers, W. (1982) Preferential codon usage in prokaryotic genes - the optimal codon anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene,* **18**, 199-209.

Gross, D.S. & Garrard, W.T. (1988) Nuclease hypersensitive sites in chromatin. *Ann Rev Biochem,* **57**, 159-197.

Grosschedl, R., Giese, K. & Pagel, J. (1994) HMG domain proteins - architectural elements in the assembly of nucleoprotein structures. *Trends Genet,* **10**, 94-100.

Grunstein, M., Durrin, L.K., Mann, R.K., Fisher-Adams, G. & Johnson, L.M. (1992) Histone: regulation of transcription in yeast. In: *Transcriptional regulation*, edited by McKnight, S. & Yamamoto, K. Cold Spring Harbor Laboratory Press, New York, p. 1295-1315.

Gualerzi, C.O., Pon, C.L., Pawlik, R.T., Canonaco, M.A., Paci, M. & Wintermeyer, W. (1986) Role of initiation factors in *Escherichia coli* translational initiation. In: *Structure, function, and genetics of ribosomes*, edited by Hardesty, B. & Kramer, G. *Springer-Verlag*, New York, p. 621-641.

Gualerzi, C.O. & Pon, C.L. (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry,* **29**, 5881-5889.

Guo, B., Odgren, P.R., Vanwijnen, A.J., Last, T.J., Nickerson, J., Penman, S., Lian, J.B., Stein, J.L. & Stein, G.S. (1995) The nuclear matrix protein Nmp-1 is the transcription factor yy1. *Proc Natl Acad Sci USA,* **92**, 10526-10530.

Gutell, R.R., Weiser, B., Woese, C.R. & Noller, H.F. (1985) Comparative anatomy of 16s-like rRNA. *Prog Nucl Acid Res Mol Biol,* **32**, 155-216.

Hager, G., Smith, C., Svaren, J. & Horz, W. (1995) Initiation of expression: remodelling genes. In: *Chromatin structure and gene expression*, edited by Elgin, S.C.R. IRL press, Oxford, p. 89-103.

Hagerman, P.J. (1988) Flexibility of DNA. *Annu Rev Biophys Biophys Chem,* **17**, 265-286.

Hagerman, P.J. & Ramadevi, V.A. (1990) Application of the method of phage T4 DNA ligase catalyzed ring closure to the study of DNA structure .1. computational analysis. *J Mol Biol,* **212**, 351-362.

Hall, J.M. & Cole, R.D. (1986) In nondividing cells, histone H1(0) is synthesized and deposited onto chromatin without accompanying phosphorylation. *Biochemistry,* **25**, 491-495.

Hall, M.N., Gabay, J., Debarbouille, M. & Schwartz, M. (1982) A role for mRNA secondary structure in the control of translation initiation. *Nature,* **295**, 616-618.

Haran, T.E., Kahn, J.D. & Crothers, D.M. (1994) Sequence elements responsible for DNA curvature. *J Mol Biol,* **244**, 135-143.

Hartman, P.G., Chapman, G.E., Moss, T. & Bradbury, E.M. (1977) Studies of the role and mode of operation of the very lysine rich histone H1 in eukaryote chromatin. *Eur J Biochem,* **77**, 45-51.

Hartz, D., Mcpheeters, D.S. & Gold, L. (1989) Selection of the initiator tRNA by *Escherichia coli* initiation factors. *Genes Devel,* **3**, 1899-1912.

Hartz, D., Mcpheeters, D.S. & Gold, L. (1991) Influence of mRNA determinants on translation initiation in *Escherichia coli. J Mol Biol,* **218**, 83-97.

Hattori, M. & Sakaki, Y. (1986) Dideoxy sequencing method using denatured plasmid templates. *Anal Biochem,* **152**, 232-238.

Hayes, J.J. (1996) Site-directed cleavage of DNA by a linker histone-Fe(II) EDTA conjugate - localization of a globular domain binding site within a nucleosome. *Biochemistry,* **35**, 11931-11937.

Hayes, J.J., Tullius, T.D. & Wolffe, A.P. (1990) The structure of DNA in a nucleosome. *Proc Natl Acad Sci USA,* **87**, 7405-7409.

Hayes, J.J., Clark, D.J. & Wolffe, A.P. (1991) Histone contributions to the structure of DNA in the nucleosome. *Proc Natl Acad Sci USA,* **88**, 6829-6833.

Hayes, J.J. & Wolffe, A.P. (1992a) The interaction of transcription factors with nucleosomal DNA. *BioEssays,* **14**, 597-603.

Hayes, J.J. & Wolffe, A.P. (1992b) Histones H2A/H2B inhibit the interaction of transcription factor IIIA with the *Xenopus borealis* somatic 5S RNA gene in a nucleosome. *Proc Natl Acad Sci USA,* **89**, 1229-1233.

Hayes, J.J. & Wolffe, A.P. (1993) Preferential and asymmetric interaction of linker histones with 5S DNA in the nucleosome. *Proc Natl Acad Sci USA,* **90**, 6415-6419.

Hayes, J.J., Pruss, D. & Wolffe, A.P. (1994) Contacts of the globular domain of histone H5 and core histones with DNA in a chromatosome. *Proc Natl Acad Sci USA,* **91**, 7817-7821.

Hayes, J.J., Kaplan, R., Ura, K., Pruss, D. & Wolffe, A. (1996) A putative DNA-binding surface in the globular domain of a linker histone is not essential for specific binding to the nucleosome. *J Biol Chem,* **271**, 25817-25822.

Hebbes, T.R., Thorne, A.W. & Crane-Robinson, C. (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J,* **7**, 1395-1402.

Hecht, A., Laroche, T., Strahlbolsinger, S., Gasser, S.M. & Grunstein, M. (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins - a molecular model for the formation of heterochromatin in yeast. *Cell,* **80**, 583-592.

Helke, A., Geisen, R.M., Vollmer, M., Sprengart, M.L. & Fuchs, E. (1993) An unstructured mRNA region and a 5′ hairpin represent important elements of the *Escherichia coli* translation initiation signal determined by using the bacteriophage T7 gene 1 translation start site. *Nucl Acids Res,* **21**, 5705-5711.

Hentze, M.W. (1997) Translation - eIF4g: a multipurpose ribosome adapter? *Science,* **275**, 500-501.

Hesse, J.E., Nickol, J.M., Lieber, M.R. & Felsenfeld, G. (1986) Regulated gene expression in transfected primary chicken erythrocytes. *Proc Natl Acad Sci USA,* **83**, 4312-4316.

Hewish, D.R. & Burgoyne, L.A. (1973) Chromatin sub structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem. Biophys. Res. Comm.*, **52**, 504-510.

Hill, C.S., Martin, S.R. & Thomas, J.O. (1989) A stable alpha-helical element in the carboxy-terminal domain of free and chromatin-bound histone H1 from sea urchin sperm. *EMBO J,* **8**, 2591-2599.

Hill, C.S., Rimmer, J.M., Green, B.N., Finch, J.T. & Thomas, J.O. (1991) Histone - DNA interactions and their modulation by phosphorylation of -ser-pro-x-lys/a. *EMBO J,* **10**, 1939-1948.

Hirano, T. & Mitchison, T.J. (1993) Topoisomerase II does not play a scaffolding role in the organization of mitotic chromosomes assembled in *Xenopus* egg extracts. *J Cell Biol,* **120**, 601-612.

Hirano, T. & Mitchison, T.J. (1994) A heterodimeric coiled-coil protein required for mitotic chromosome condensation *in vitro*. *Cell,* **79**, 449-458.

Hiraoka, Y., Minden, J.S., Swedlow, J.R., Sedat, J.W. & Agard, D.A. (1989) Focal points for chromosome condensation and decondensation revealed by 3-dimensional *in vivo* time lapse microscopy. *Nature,* **342**, 293-296.

Hirschhorn, J.N., Brown, S.A., Clark, C.D. & Winston, F. (1992) Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes Dev,* **6**, 2288-2298.

Holt, C.L. & May, G.S. (1996) An extragenic suppressor of the mitosis defective bimd6 mutation of *Aspergillus nidulans* codes for a chromosome scaffold protein. *Genetics,* **142**, 777-787.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P. & Bradbury, E.M. (1993) Studies of the DNA-binding properties of histone H4 amino terminus – thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 tail to DNA. *J Biol Chem,* **268**, 305-314.

Horiuchi, J., Silverman, N., Marcus, G.A. & Guarente, L. (1995) Ada3, a putative transcriptional adapter, consists of 2 separable domains and interacts with Ada2 and Gcn5 in a trimeric complex. *Mol Cell Biol,* **15**, 1203-1209.

Hui, A., Hayflick, J., Dinkelspiel, K. & Deboer, H.A. (1984) Mutagenesis of the 3 bases preceding the start codon of the β-galactosidase mRNA and its effect on translation in *Escherichia coli. EMBO J,* **3**, 623-629.

Hui, A. & De Boer, H.A. (1987) Specialized ribosome system - preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli. Proc Natl Acad Sci USA,* **84**, 4762-4766.

Imbalzano, A.N., Kwon, H., Green, M.R. & Kingston, R.E. (1994) Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature,* **370**, 481-485.

Inouye, S. & Tsuji, F.I. (1994) *Aequorea* green fluorescent protein - expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Lett,* **341**, 277-280.

Ioshikhes, I., Bolshoy, A. & Trifonov, E.N. (1992) Preferred positions of AA-dinucleotides and TT-dinucleotides in aligned nucleosomal DNA-sequences. *J Biomol Struc Dynam,* **9**, 1111-1117.

Ioshikhes, I., Bolshoy, A., Derenshteyn, K., Borodovsky, M. & Trifonov, E.N. (1996) Nucleosome DNA sequence pattern revealed by multiple alignment of experimentally mapped sequences. *J Mol Biol,* **262**, 129-139.

Ito, K., Kawakami, K. & Nakamura, Y. (1993) Multiple control of *Escherichia coli* lysyl-tRNA synthetase expression involves a transcriptional repressor and translational enhancer element. *Proc Natl Acad Sci USA,* **90**, 302-306.

Izaurralde, E., Kas, E. & Laemmli, U.K. (1989) Highly preferential nucleation of histone H1 assembly on scaffold associated regions. *J Mol Biol,* **210**, 573-585.

Jackson, J.R. & Benyajati, C. (1993) DNA-histone interactions are sufficient to position a single nucleosome juxtaposing *Drosophila* Adh adult enhancer and distal promoter. *Nucl Acids Res,* **21**, 957-967.

Jacob, W.F., Santer, M. & Dahlberg, A.E. (1987) A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proc Natl Acad Sci USA,* **84**, 4757-4761.

Johnston, M. & Dover, J. (1987) Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA-binding domain. *Proc Natl Acad Sci USA,* **84**, 2401-2405.

Juan, L.J., Utley, R.T., Vignali, M., Bohm, L. & Workman, J.L. (1997) H1-mediated repression of transcription factor binding to a stably positioned nucleosome. *J Biol Chem,* **272**, 3635-3640.

Kamakaka, R.T. & Thomas, J.O. (1990) Chromatin structure of transcriptionally competent and repressed genes. *EMBO J,* **9**, 3997-4006.

Kandolf, H. (1994) The H1a histone variant is an *in vivo* repressor of oocyte type 5S gene transcription in *Xenopus laevis* embryos. *Proc Natl Acad Sci USA,* **91**, 7257-7261.

Kas, E., Izaurralde, E. & Laemmli, U.K. (1989) Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin - the role of oligo(dA)•oligo(dT) tracts. *J Mol Biol,* **210**, 587-599.

Kaufman, P.D. & Botchan, M.R. (1994) Assembly of nucleosomes: do multiple assembly factors mean multiple mechanism? *Curr Opin Genet Dev,* **4**, 229-235.

Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F. & Grunstein, M. (1988) Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell,* **55**, 27-39.

Kefalas, P., Gray, F.C. & Allan, J. (1988) Precise nucleosome positioning in the promoter of the chicken β^Aglobin gene. *Nucl Acids Res,* **16**, 501-517.

Khadake, J.R. & Rao, M.R.S. (1997) Condensation of DNA and chromatin by an SPKK-containing octapeptide repeat motif present in the C-terminus of histone H1. *Biochemistry,* **36**, 1041-1051.

Klimyuk, V.I. & Karpenchuk, K.G. (1988) Length of the nucleosomal repeat of DNA and variants of the core histones of higher plants. *Biochemistry-Ussr,* **53**, 389-396.

Klug, A., Rhodes, D., Smith, J., Finch, J.T. & Thomas, J.O. (1980) A low resolution structure for the histone core of the nucleosome. *Nature,* **287**, 509-516.

Knezetic, J.A. & Luse, D.S. (1986) The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II *in vitro*. *Cell,* **45**, 95-104.

Koo, H.S., Wu, H.M. & Crothers, D.M. (1986) DNA bending at A•T tracts. *Nature,* **320**, 501-506.

Koo, H.S. & Crothers, D.M. (1988) Calibration of DNA curvature and a unified description of sequence- directed bending. *Proc Natl Acad Sci USA,* **85**, 1763-1767.

Kornberg, R.D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science,* **184**, 868-871.

Kornberg, R.D. & Thomas, J.O. (1974) Chromatin structure: oligomers of histones. *Science,* **184**, 865-868.

Kornberg, R.D. & Stryer, L. (1988) Statistical distributions of nucleosomes - nonrandom locations by a stochastic mechanism. *Nucl Acids Res,* **16**, 6677-6690.

Kozak, M. (1989) The scanning model for translation - an update. *J Cell Biol,* **108**, 229-241.

Kralovics, R., Fajkus, J., Kovarik, A. & Bezdek, M. (1995) DNA curvature of the tobacco grs repetitive sequence family and its relation to nucleosome positioning. *J Biomol Struc Dynam,* **12**, 1103-1119.

Kropp, B., Leoni, L., Sampaolese, B. & Savino, M. (1995) Influence of DNA superstructural features and histone amino terminal domains on nucleosome positioning. *FEBS Lett,* **364**, 17-22.

Kruger, W. & Herskowitz, I. (1991) A negative regulator of H^0 transcription, Sin1 (Spt2), is a nonspecific DNA-binding protein related to HMG1. *Mol Cell Biol,* **11**, 4135-4146.

Kruger, W., Peterson, C.L., Sil, A., Coburn, C., Arents, G., Moudrianakis, E.N. & Herskowitz, I. (1995) Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes Dev,* **9**, 2770-2779.

Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y. & Allis, C.D. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature,* **383**, 269-272.

Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E. & Green, M.R. (1994) Nucleosome disruption and enhancement of activator binding by a human SWI/SNFcomplex. *Nature,* **370**, 477-481.

Lamphear, B.J., Kirchweger, R., Skern, T. & Rhoads, R.E. (1995) Mapping of functional domains in eukaryotic protein synthesis initiation-factor 4g (eIF4g) with picornaviral proteases -implications for cap-dependent and cap-independent translational initiation. *J Biol Chem,* **270**, 21975-21983.

Laughrea, M. & Tam, J. (1989) Ribosomal protein S1 and initiation factor IF3 do not promote the ribosomal binding of ~19-nucleotide long mRNA and mRNA models. *Biochem Cell Biol,* **67**, 812-817.

Laurent, B.C., Treich, I. & Carlson, M. (1993) The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. *Genes Dev,* **7**, 583-591.

Lea, M.A. (1987) Relationship of $H1^0$ histone to differentiation and cancer. *Cancer Biochem Biophys,* **9**, 199-209.

Lee, D.Y., Hayes, J.J., Pruss, D. & Wolffe, A.P. (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell,* **72**, 73-84.

Lennard, A.C. & Thomas, J.O. (1985) The arrangement of H5 molecules in extended and condensed chicken erythrocyte chromatin. *EMBO J,* **4**, 3455-3462.

Leuba, S.H., Yang, G.L., Robert, C., Samori, B., van Holde, K., Zlatanova, J. & Bustamante, C. (1994) 3-dimensional structure of extended chromatin fibers as revealed by tapping mode scanning force microscopy. *Proc Natl Acad Sci USA,* **91**, 11621-11625.

Leuba, S.H., Zlatanova, J. & van Holde, K. (1993) On the location of histone H1 and histone H5 in the chromatin fiber - studies with immobilized trypsin and chymotrypsin. *J Mol Biol,* **229**, 917-929.

Lewis, E.A. & Reams, R.R. (1983) Histone H1 – ultra centrifugation studies of the effects of ionic strength and denaturants on the solution conformation. *Arch Biochem Biophys,* **223**, 185-192.

Li, B.Y., Adams, C.C. & Workman, J.L. (1994) Nucleosome binding by the constitutive transcription factor Sp1. *J Biol Chem,* **269**, 7756-7763.

Li, Q. & Wrange, O. (1993) Translational positioning of a nucleosomal glucocorticoid response element modulates glucocorticoid receptor affinity. *Genes Dev,* **7**, 2471-2482.

Li, W., Nagaraja, S., Delcuve, G.P., Hendzel, M.J. & Davie, J.R. (1993) Effects of histone acetylation, ubiquitination and variants on nucleosome stability. *Biochem Journal* **296**, 737-744.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L. & Ward, D.C. (1988) Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* supression hybridization using recombinant DNA libraries. *Hum Genet,* **80**, 224-234.

Liiv, A., Tenson, T. & Remme, J. (1996) Analysis of the ribosome large subunit assembly and 23S rRNA stability *in vivo. J Mol Biol,* **263**, 396-410.

Lorch, Y., Lapointe, J.W. & Kornberg, R.D. (1987) Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell,* **49**, 203-210.

Losa, R. & Brown, D.D. (1987) A bacteriophage RNA polymerase transcribes *in vitro* through a nucleosome core without displacing it. *Cell,* **50**, 801-808.

Lowman, H. & Bina, M. (1990) Correlation between dinucleotide periodicities and nucleosome positioning on mouse satellite DNA. *Biopolymers,* **30**, 861-876.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* **389**, 251-260.

Lutter, L.C. (1979) Precise location of DNase I cutting sites in the nucleosome core determined by high resolution gel electrophoresis. *Nucleic Acids Res,* **6**, 41-56.

Makarov, V., Dimitrov, S., Smirnov, V. & Pashev, I. (1985) A triple helix model for the structure of chromatin fiber. *FEBS Lett,* **181**, 357-361.

Makrides, S.C. (1996) Strategies for achieving high level expression of gene in *Escherichia coli. Microbiol Rev,* **60**, 512-538.

Maman, J.D., Yager, T.D. & Allan, J. (1994) Self-association of the globular domain of histone H5. *Biochemistry,* **33**, 1300-1310.

Marcus, G.A., Silverman, N., Berger, S.L., Horiuchi, J. & Guarente, L. (1994) Functional similarity and physical association between Gcn5 and Ada2 - putative transcriptional adapters. *EMBO J,* **13**, 4807-4815.

Marino, M.H. (1989) Expression system for heterologous protein production. *Bio Pharm,* **2**, 18-33.

Matallana, E., Franco, L. & Perezortin, J.E. (1992) Chromatin structure of the yeast suc2 promoter in regulatory mutants. *Mol Gen Genet,* **231**, 395-400.

Mathog, D., Hochstrasser, M., Gruenbaum, Y., Saumweber, H. & Sedat, J. (1984) Characteristic folding pattern of polytene chromosomes in *Drosophila salivary* gland nuclei. *Nature,* **308**, 414-421.

McCarthy, J.E.G., Schairer, H.U. & Sebald, W. (1985) Translational initiation frequency of ATP genes from *Escherichia coli* - identification of an intercistronic sequence that enhances translation. *EMBO J,* **4**, 519-526.

McCarthy, J.E.G. & Gaulerzi, C. (1990) Translational control of prokaryotic gene expression. *Trends Genet,* **6**, 78-85.

McCarthy, J.E.G. & Brimacombe, R. (1994) Prokaryotic translation: the interactive pathway leading to initiation. *Trends Genet,* **10**, 402-407.

McGhee, J.D. & Felsenfeld, G. (1980) Nucleosome structure. *Ann Rev Biochem,* **49**, 1115-1156.

McGhee, J.D., Rau, D.C., Charney, E. & Felsenfeld.G. (1980) Orientation of the nucleosomes within the higher order structure of chromatin. *Cell,* **22**, 87-96.

McGhee, J.D., Wood, W.I., Dolan, M., Engel, J.D. & Felsenfeld, G. (1981) A 200 base pair region at the 5′ end of the chicken adult β globin gene is accessible to nuclease digestion. *Cell,* **27**, 45-55.

McGhee, J.D., Nickol, J.M., Felsenfeld, G. & Rau, D.C. (1983) Higher order structure of chromatin - orientation of nucleosomes within the 30 nm chromatin solenoid is independent of species and spacer length. *Cell,* **33**, 831-841.

McNamara, P.T., Bolshoy, A., Trifonov, E.N. & Harrington, R.E. (1990) Sequence-dependent kinks induced in curved DNA. *J Biomol Struc Dynam,* **8**, 529-538.

Meersseman, G., Pennings, S. & Bradbury, E.M. (1991) Chromatosome positioning on assembled long chromatin linker histones affect nucleosome placement on 5S rDNA. *J Mol Biol,* **220**, 89-100.

Meersseman, G., Pennings, S. & Bradbury, E.M. (1992) Mobile nucleosomes - a general behavior. *EMBO J,* **11**, 2951-2959.

Merrick, W.C. & Hershey, J.W.B. (1996) The pathway and mechanism of eukaryotic protein synthesis. In: *Translational control*, edited by Hershey, J.W.B., Mathews, M.B. & Sonenberg, N. Cold Spring Harbor Laboratory Press, New York, p. 31-69.

Mirkovitch, J., Mirault, M.E. & Laemmli, U.K. (1984) Organization of the higher order chromatin loop - specific DNA attachment sites on nuclear scaffold. *Cell,* **39**, 223-232.

Mirkovitch, J., Gasser, S.M. & Laemmli, U.K. (1988) Scaffold attachment of DNA loops in metaphase chromosomes. *J Mol Biol,* **200**, 101-109.

Mirzabekov, A.D., Shick, V.V., Belyavsky, A.V. & Bavykin, S.G. (1978) Primary organization of nucleosome core particle in chromatin: sequence of histone arrangement along DNA. *Proc Natl Acad Sci USA,* **75**, 4184-4189.

Moqtaderi, Z., Yale, J.D., Struhl, K. & Buratowski, S. (1996) Yeast homologs of higher eukaryotic TFIID subunits. *Proc Natl Acad Sci USA,* **93**, 14654-14658.

Morise, J.G., Shimomura, O., Johnson, F.H. & Winant, J. (1974) Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry,* **13**, 2656-2662.

Moudrianakis, E.N. & Arents, G. (1993) Structure of the histone octamer core of the nucleosome and its potential interactions with DNA. *Cold Spring Harbor Symp Quant Biol,* **58**, 273-279.

Movva, N.R., Nakamura, K. & Inouye, M. (1980) Gene structure of the OmpA protein, a major surface protein of *Escherichia coli* required for cell-cell interaction. *J Mol Biol,* **143**, 317-328.

Mueller, R.D., Yasuda, H. & Bradbury, E.M. (1985a) Phosphorylation of histone H1 through the cell cycle of Physarum polycephalum: 24 sites of phosphorylation at metaphase. *J Biol Chem,* **260**, 5081-5086.

Mueller, R.D., Yasuda, H., Hatch, C.L., Bonner, W.M. & Bradbury, E.M. (1985b) Identification of ubiquitinated histones 2A and 2B in *Physarum polycephalum.* Disappearance of these proteins at metaphase and reappearance at anaphase. *J Biol Chem,* **260**, 5147-5153.

Muyldermans, S. & Travers, A.A. (1994) DNA sequence organization in chromatosomes. *J Mol Biol,* **235**, 855-870.

Nacheva, G.A., Preobrazhenskaya, O.V., Melnikova, A.F. & Karpov, V.L. (1989) Structural changes in chromatin of the *Drosophila* heat shock protein 70 gene during transcription. *Mol Biol.* **23**, 695-703.

Nagai, H., Yuzawa, H. & Yura, T. (1991) Interplay of 2 *cis*-acting mRNA regions in translational control of sigma 32 synthesis during the heat shock response of *Escherichia coli. Proc Natl Acad Sci USA.* **88**, 10515-10519.

Neubauer, B., Linxweiler, W. & Horz, W. (1986) DNA engineering shows that nucleosome phasing on the African green monkey alpha-satellite is the result of multiple additive histone-DNA interactions. *J Mol Biol,* **190**, 639-645.

Nicaud, J.M., Mackman, N. & Holland, I.B. (1986) Current status of secretion of foreign proteins by microorganisms. *J Biotechnol,* **3**, 255-270.

Nickel, B.E., Allis, C.D. & Davie, J.R. (1989) Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochemistry,* **28**, 958-963.

Nickol, J., Behe, M. & Felsenfeld, G. (1982) Effect of the B-Z transition in poly(dG-m^5dC).poly(dG-m^5dC) on nucleosome formation. *Proc Natl Acad Sci USA - Biological Sciences,* **79**, 1771-1775.

Nissen, M.S. & Reeves, R. (1995) Changes in superhelicity are introduced into closed circular DNA by binding of high-mobility group protein I/Y. *J Biol Chem.* **270**, 4355-4360.

Nobile, C., Nickol, J. & Martin, R.G. (1986) Nucleosome phasing on a DNA fragment from the replication origin of simian virus 40 and rephasing upon cruciform formation of the DNA. *Mol Cell Biol,* **6**, 2916-2922.

Noll, M. (1974) Internal structure of the chromatin subunit. *Nucleic Acids Res,* **1**, 1573-1578.

Noll, M. & Kornberg, R.D. (1977) Action of micrococcal nuclease on chromatin and the location of histone H1. *J Mol Biol,* **109**, 393-404.

Norton, V.G., Imai, B.S., Yau, P. & Bradbury, E.M. (1989) Histone acetylation reduces nucleosome core particle linking number change. *Cell,* **57**, 449-457.

Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. & Nakatani, Y. (1996) The transcriptional coactivators p300 and cbp are histone acetyltransferases. *Cell,* **87**, 953-959.

Ohalloran, T.V., Lippard, S.J., Richmond, T.J. & Klug, A. (1987) Multiple heavy-atom reagents for macromolecular X-ray structure determination - application to the nucleosome core particle. *J Mol Biol,* **194**, 705-712.

Ohsumi, K., Katagiri, C. & Kishimoto, T. (1993) Chromosome condensation in *Xenopus* mitotic extracts without histone H1. *Science,* **262**, 2033-2035.

Olins, P.O. & Rangwala, S.H. (1989) A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the lacz gene in *Escherichia coli. J Biol Chem,* **264**, 16973-16976.

Olins, P.O. & Lee, S.C. (1993) Recent advances in heterologous gene expression in *Escherichia coli. Curr Opin Biotechnol,* **4**, 520-525.

Oliviero, S. & Struhl, K. (1991) Synergistic transcriptional enhancement does not depend on the number of acidic activation domains bound to the promoter. *Proc Natl Acad Sci USA,* **88**, 224-228.

Olson, W.K., Marky, N.L., Jernigan, R.L. & Zhurkin, V.B. (1993) Influence of fluctuations on DNA curvature - a comparison of flexible and static wedge models of intrinsically bent DNA. *J Mol Biol,* **232**, 530-551.

Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. & Sigler, P.B. (1988) Crystal-structure of Trp repressor operator complex at atomic resolution. *Nature,* **335**, 321-329.

Oudet, P., Gross-Bellard, M. & Chambon, P. (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell,* **4**, 281-300.

Owen-Hughes, T., Utley, R.T., Cote, J., Peterson, C.L. & Workman, J.L. (1996) Persistent site specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. *Science,* **273**, 513-516.

Owen-Hughes, T. & Workman, J.L. (1996) Remodeling the chromatin structure of a nucleosome array by transcription factor targeted *trans* displacement of histones. *EMBO J,* **15**, 4702-4712.

Paranjape, S.M., Kamakaka, R.T. & Kadonaga, J.T. (1994) Role of chromatin structure in the regulation of transcription by RNA polymerase II. *Annu Rev Biochem,* **63**, 265-297.

Park, Y.W. & Breslauer, K.J. (1991) A spectroscopic and calorimetric study of the melting behaviors of a bent and a normal DNA duplex - $[d(GA_4T_4C)]_2$ versus $[d(GT_4A_4C)]_2$. *Proc Natl Acad Sci USA,* **88**, 1551-1555.

Paulson, J.R. & Laemmli, U.K. (1977) The structure of histone-depleted metaphase chromosomes. *Cell,* **12**, 817

Pazin, M.J., Kamakaka, R.T. & Kadonaga, J.T. (1994) ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. *Science,* **266**, 2007-2011.

Pederson, D.S., Thoma, F. & Simpson, R.T. (1986) Core particle, fiber, and transcriptionally active chromatin structure. *Annu Rev Cell Biol,* **2**, 117-147.

Pehrson, J.R. (1989) Thymine dimer formation as a probe of the path of DNA in and between nucleosomes in intact chromatin. *Proc Natl Acad Sci USA,* **86**, 9149-9153.

Pelka, H. & Schulman, L.H. (1986) Study of the interaction of *Escherichia coli* $tRNA^{met}$ synthetase with $tRNA^{fmet}$ using chemical and enzymatic probes. *Biochemistry,* **25**, 4450-4456.

Pennings, S., Muyldermans, S., Meersseman, G. & Wyns, L. (1989) Formation, stability and core histone positioning of nucleosomes reassembled on bent and other nucleosome-derived DNA. *J Mol Biol,* **207**, 183-192.

Pennings, S., Meersseman, G. & Bradbury, E.M. (1991) Mobility of positioned nucleosomes on 5S rDNA. *J Mol Biol,* **220**, 101-110.

Pennings, S., Meersseman, G. & Bradbury, E.M. (1994) Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proc Natl Acad Sci USA,* **91**, 10275-10279.

Perezortin, J.E., Matallana, E. & Franco, L. (1989) Chromatin structure of yeast genes. *Yeast,* **5**, 219-238.

Perlmann, T. & Wrange, O. (1988) Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome. *EMBO J,* **7**, 3073-3079.

Perry, C.A., Dadd, C.A., Allis, C.D. & Annunziato, A.T. (1993) Analysis of nucleosome assembly and histone exchange using antibodies specific for acetylated H4. *Biochemistry,* **32**, 13605-13614.

Pilipenko, E.V., Gmyl, A.P., Maslova, S.V., Svitkin, Y.V., Sinyakov, A.N. & Agol, V.I. (1992) Prokaryotic-like *cis* elements in the cap-independent internal initiation of translation on picornavirus RNA. *Cell,* **68**, 119-131.

Pina, B., Bruggemeier, U. & Beato, M. (1990) Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. *Cell,* **60**, 719-731.

Polach, K.J. & Widom, J. (1995) Mechanism of protein access to specific DNA-sequences in chromatin - a dynamic equilibrium model for gene regulation. *J Mol Biol,* **254**, 130-149.

Powers, T. & Noller, H.F. (1995) A temperature-dependent conformational rearrangement in the ribosomal protein S4 center dot 16S rRNA complex. *J Biol Chem,* **270**, 1238-1242.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. & Cormier, M.J. (1992) Primary structure of the *Aequorea victoria* green fluorescent protein. *Gene* **111**, 229-233.

Pratt, K. & Hattman, S. (1983) Nucleosome phasing in *Tetrahymena* macronuclei. *J Protozoology,* **30**, 592-598.

Pruss, D. & Wolffe, A.P. (1993) Histone DNA contacts in a nucleosome core containing a *Xenopus* 5S rRNA gene. *Biochemistry,* **32**, 6810-6814.

Pruss, D., Hayes, J.J. & Wolffe, A.P. (1995) Nucleosomal anatomy - where are the histones. *BioEssays,* **17**, 161-170.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E.N. & Wolffe, A.P. (1996) An asymmetric model for the nucleosome - a binding site for linker histones inside the DNA gyres. *Science,* **274**, 614-617.

Ptashne, M. (1986) Gene regulation by proteins acting nearby and at a distance. *Nature,* **322**, 697-701.

Quinn, J., Fyrberg, A.M., Ganster, R.W., Schmidt, M.C. & Peterson, C.L. (1996) DNA-binding properties of the yeast SWI/SNF complex. *Nature,* **379**, 844-847.

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. & Sweet, R.M. (1993) Crystal-structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature,* **362**, 219-223.

Ramakrishnan, V. (1995) The histone fold - evolutionary questions. *Proc Natl Acad Sci USA,* **92**, 11328-11330.

Ramesh, V., De, A. & Nagaraja, V. (1994) Engineering hyperexpression of bacteriophage mu C protein by removal of secondary structure at the translation initiation region. *Protein Eng,* **7**, 1053-1057.

Ramsay, N. (1986) Deletion analysis of a DNA sequence that positions itself precisely on the nucleosome core. *J Mol Biol,* **189**, 179-188.

Reese, J.C., Apone, L., Walker, S.S., Griffin, L.A. & Green, M.R. (1994) Yeast TAF(II)s in a multisubunit complex required for activated transcription. *Nature,* **371**, 523-527.

Reeves, R. & Nissen, M.S. (1993) Interaction of high mobility group I(Y) nonhistone proteins with nucleosome core particles. *J Biol Chem,* **268**, 21137-21146.

Reik, A., Schutz, G. & Stewart, A.F. (1991) Glucocorticoids are required for establishment and maintenance of an alteration in chromatin structure - induction leads to a reversible disruption of nucleosomes over an enhancer. *EMBO J,* **10**, 2569-2576.

Renz, M. & Day, L.A. (1976) Transition from noncooperative of cooperative and selective binding of histone H1 to DNA. *Biochemistry,* **15**, 3220-3228.

Renz, M., Nehls, P. & Hozier, J. (1977) Involvement of histone H1 in the organization of the chromosome fiber. *Proc Natl Acad Sci USA,* **74**, 1879-1883.

Resch, A., Tedin, K., Grundling, A., Mundlein, A. & Blasi, U. (1996) Downstream box-anti-downstream box interactions are dispensable for translation initiation of leaderless mRNAs. *EMBO J,* **15**, 4740-4748.

Rhodes, D. (1985) Structural analysis of a triple complex between the histone octamer, a *Xenopus* gene for 5S RNA and transcription factor III A. *EMBO J,* **4**, 3473-3482.

Richard-Foy, H. & Hager, G.L. (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *EMBO J,* **6**, 2321-2328.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. & Klug, A. (1984) Structure of the nucleosome core particle at 7Å resolution. *Nature,* **311**, 532-537.

Ridsdale, J.A., Hendzel, M.J., Delcuve, G.P. & Davie, J.R. (1990) Histone acetylation alters the capacity of the H1 histones to condense transcriptionally active competent chromatin. *J Biol Chem,* **265**, 5150-5156.

Ring, D. & Cole, R.D. (1983) Close contacts between H1-histone molecules in nuclei. *J Biol Chem,* **258**, 5361-5364.

Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G.D. & Gold, L. (1992) Translation initiation in *Escherichia coli* - sequences within the ribosome binding site. *Mol Microbiol,* **6**, 1219-1229.

Roth, S.Y., Schulman, I.G., Richman, R., Cook, R.G. & Allis, C.D. (1988) Characterization of phosphorylation sites in histone H1 in the amitotic macronucleus of tetrahymena during different physiological states. *J Cell Biol,* **107**, 2473-2482.

Roth, S.Y., Dean, A. & Simpson, R.T. (1990) Yeast α-2 repressor positions nucleosomes in trp1 ars1 chromatin. *Mol Cell Biol,* **10**, 2247-2260.

Roth, S.Y. & Allis, C.D. (1992) Chromatin condensation - does histone H1 dephosphorylation play a role. *Trends Biochem Sci,* **17**, 93-98.

Ruppert, S. & Tjian, R. (1995) Human TAF(pi)250 interacts with rap74 - implications for RNA polymerase II initiation. *Genes Dev,* **9**, 2747-2755.

Russanova, V.R., Dimitrov, S.I., Makarov, V.L. & Pashev, I.G. (1987) Accessibility of the globular domain of histone H1 and histone H5 to antibodies upon folding of chromatin. *Eur J Biochem,* **167**, 321-326.

Sambrook, J., Fritsh, E.F. & Maniatis, T. (1989) Molecular cloning. A laboratory manual. 3rd ed, Cold Spring Harbor Laboratory Press, New York,

Sandeen, G., Wood, W.I. & Felsenfeld, G. (1980) The interaction of high mobility group proteins HMG14 and 17 with nucleosomes. *Nucleic Acids Res,* **8**, 3757-3778.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA,* **74**, 5463-5467.

Satchwell, S.C., Drew, H.R. & Travers, A.A. (1986) Sequence periodicities in chicken nucleosome core DNA. *J Mol Biol,* **191**, 659-675.

Satchwell, S.C. & Travers, A.A. (1989) Asymmetry and polarity of nucleosomes in chicken erythrocyte chromatin. *EMBO J,* **8**, 229-238.

Sauer, F., Wassarman, D.A., Rubin, G.M. & Tjian, R. (1996) TAF(II)s mediate activation of transcription in the drosophila embryo. *Cell,* **87**, 1271-1284.

Scherer, G.F.E., Walkinshaw, M.D., Arnott, S. & Morre, J.D. (1980) The ribosome binding sites recognized by *E.coli* ribosomes have regions with signal character in both the leader and protein coding segments. *Nucleic Acids Res,* **8**, 3895-3907.

Schild, C., Claret, F.X., Wahli, W. & Wolffe, A.P. (1993) A nucleosome dependent static loop potentiates estrogen-regulated transcription from the *Xenopus vitellogenin* b1 promoter invitro. *EMBO J,* **12**, 423-433.

Schneider, T., Stormo, G.D., Gold, L. & Ehrenfeuchs, A. (1986) Information content of binding sites on nucleotide sequences. *J Mol Biol,* **188**, 415-431.

Shatzman, A.R. (1995) Expression systems - editorial overview. *Curr Opin Biotechnol,* **6**, 491-493.

Shean, C.S. & Gottesman, M.E. (1992) Translation of the prophage-λ cI transcript. *Cell,* **70**, 513-522.

Shen, X.T., Yu, L.L., Weir, J.W. & Gorovsky, M.A. (1995) Linker histones are not essential and affect chromatin condensation *in vivo*. *Cell,* **82**, 47-56.

Shen, X.T. & Gorovsky, M.A. (1996) Linker histone H1 regulates specific gene-expression but not global transcription *in vivo*. *Cell,* **86**, 475-483.

Shick, V.V., Belyavsky, A.G., Bavykin, S.G. & Mirzabekov, A.D. (1980) Primary organization of the nucleosome core particles. Sequential arrangement of histones along DNA. *J Mol Biol,* **139**, 491-517.

Shimizu, M., Roth, S.Y., Szentgyorgyi, C. & Simpson, R.T. (1991) Nucleosomes are positioned with base pair precision adjacent to the α-2 operator in *Saccharomyces cerevisiae*. *EMBO J,* **10**, 3033-3041.

Shine, J. & Dalgarno, L. (1974) The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA,* **71**, 1342-1346.

Shpigelman, E.S., Trifonov, E.N. & Bolshoy, A. (1993) Curvature - software for the analysis of curved DNA. *CABIOS,* **9**, 435-440.

Shrader, T.E. & Crothers, D.M. (1989) Artificial nucleosome positioning sequences. *Proc Natl Acad Sci USA,* **86**, 7418-7422.

Silverman, J.A. & Hill, B.A. (1995) Characterization of the basal and carcinogen regulatory elements of the rat mdr1b promoter. *Mol Carcinogenesis,* **13**, 50-59.

Simpson, R.T. (1978) Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry,* **17**, 5524-5531.

Simpson, R.T. (1991) Nucleosome positioning - occurrence, mechanisms, and functional consequences. *Prog Nucl Acid Res Mol Biol,* **40**, 143-184.

Singer, B.S., Gold, L., Shinedling, S.T., Colkitt, M., Hunter, L.R., Pribnow, D. & Nelson, M.A. (1981) Analysis *in vivo* of translational mutants of the rIIB cistron of bacteriophage T4. *J Mol Biol,* **149**, 405-432.

Sivolob, A.V. & Khrapunov, S.N. (1995) Translational positioning of nucleosomes on DNA - the role of sequence-dependent isotropic DNA bending stiffness. *J Mol Biol,* **247**, 918-931.

Spanjaard, R.A., Vandijk, M.C.M., Turion, A.J. & Vanduin, J. (1989) Expression of the rat interferon-α-1 gene in *Escherichia coli* controlled by the secondary structure of the translation initiation region. *Gene,* **80**, 345-351.

Spiker, S. (1982) Histone variants in plants - evidence for primary structure variants differing in molecular weight. *J Biol Chem,* **257**, 4250-4255.

Sprengart, M.L., Fatscher, H.P. & Fuchs, E. (1990) The initiation of translation in *Escherichia coli* - apparent base- pairing between the 16S rRNA and downstream sequences of the mRNA. *Nucl Acids Res,* **18**, 1719-1723.

Sprengart, M.L., Fuchs, E. & Porter, A.G. (1996) The downstream box - an efficient and independent translation initiation signal in *Escherichia coli. EMBO J,* **15**, 665-674.

Sprous, D., Zacharias, W., Wood, Z.A. & Harvey, S.C. (1995) Dehydrating agents sharply reduce curvature in DNAs containing A-tracts. *Nucl Acids Res,* **23**, 1816-1821.

Srinivasan, A.R., Torres, R., Clark, W. & Olson, W.K. (1987) Base sequence effects in double helical DNA .1. potential energy estimates of local base morphology. *J Biomol Struc Dynam,* **5**, 459

Staynov, D.Z. (1983) Possible nucleosome arrangements in the higher order structure of chromatin. *Intl J Biol Macromolec,* **5**, 3-9.

Staynov, D.Z. & CraneRobinson, C. (1988) Footprinting of linker histones H5 and histones H1 on the nucleosome. *EMBO J* **7**, 3685-3691.

Stefanovsky, V.Y., Dimitrov, S.I., Russanova, V.R., Angelov, D. & Pashev, I.G. (1989) Laser-induced crosslinking of histones to DNA in chromatin and core particles - implications in studying histone-DNA interactions. *Nucl Acids Res,* **17**, 10069-10081.

Steger, D.J. & Workman, J.L. (1996) Remodeling chromatin structures for transcription: what happens to the histones? *BioEssays,* **18**, 875-884.

Stein, A. & Kunzler, P. (1983) Histone H5 can correctly align randomly arranged nucleosomes in a defined *in vitro* system. *Nature,* **302**, 548-550.

Stein, A. & Bina, M. (1984) A model chromatin assembly system - factors affecting nucleosome spacing. *J Mol Biol,* **178**, 341-363.

Steitz, J.A. (1975) Ribosome recognition of initiator regions in the RNA bacteriophage genome. In: *RNA phage*, edited by Zinder, N.D. Cold Spring Harbor Laboratory Press, New York, p. 319-352.

Steitz, J.A. & Jakes, K. (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3′ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc Natl Acad Sci USA,* **72**, 4734-4738.

Steitz, T.A. (1990) Structural studies of protein nucleic acid interaction - the sources of sequence-specific binding. *Quart Rev Biophys,* **23**, 205

Stormo, G.D., Schneider, T. & Gold, L.M. (1982) Characterization of translational initiation sites in *E.coli*. *Nucleic Acids Res,* **10**, 2971-2996.

Stormo, G.D. (1986) Translation initiation. In: *Maximizing gene expression*, edited by Reznikoff, W. & Gold, L. Butterworth, Boston, p. 195-224.

Struck, M.M., Klug, A. & Richmond, T.J. (1992) Comparison of X-ray structures of the nucleosome core particle in 2 different hydration states. *J Mol Biol,* **224**, 253-264.

Studitsky, V.M., Clark, D.J. & Felsenfeld, G. (1995) Overcoming a nucleosomal barrier to transcription. *Cell,* **83**, 19-27.

Sugarmen, B.J., Dodgson, J.B. & Engel, J.D. (1983) Genomic organization, DNA sequence, and expression of chicken embryonic histone genes. *J Biol Chem,* **258**, 9005-9016.

Suzuki, M. (1988) The DNA-binding unit in sea-urchin spermatogenous histones .2. the position of the DNA-binding unit in the nucleosome structure. *Proc Japan Acad Series B-Physical And Biological Sciences,* **64**, 33-36.

Suzuki, M. (1989) SPKK, a new nucleic acid binding unit of protein found in histone. *EMBO J,* **8**, 797-804.

Suzuki, M., Sohma, H., Yazawa, M., Yagi, K. & Ebashi, S. (1990) Histone H1 kinase specific to the SPKK motif. *J Biochem,* **108**, 356-364.

Suzuki, M., Gerstein, M. & Johnson, T. (1993) An NMR-study on the DNA-binding SPKK motif and a model for its interaction with DNA. *Protein Eng,* **6**, 565-574.

Svaren, J. & Horz, W. (1996) Regulation of gene expression by nucleosomes. *Curr Opin Genet Dev,* **6**, 164-170.

Svaren, J. & Horz, W. (1997) Transcription factors vs nucleosomes: regulation of the Pho5 promoter in yeast. *Trends Biochem Sci,* **22**, 93-97.

Swedlow, J.R., Sedat, J.W. & Agard, D.A. (1993) Multiple chromosomal populations of topoisomerase-II detected *in vivo* by time-lapse, 3-dimensional wide-field microscopy. *Cell,* **73**, 97-108.

Taunton, J., Hassig, C.A. & Schreiber, S.L. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator rpd3p. *Science,* **272**, 408-411.

Taylor, I.C.A., Workman, J.L., Schuetz, T.J. & Kingston, R.E. (1991) Facilitated binding of Gal4 and heat shock factor to nucleosomal templates - differential function of DNA-binding domains. *Genes Dev,* **5**, 1285-1298.

Tessier, L.H., Sondermeyer, P., Faure, T., Dreyer, D., Benavente, A., Villeval, D., Courtney, M. & Lecocq, J.P. (1984) The influence of mRNA primary and secondary structure on human IFN-γ gene expression in *E.coli. Nucleic Acids Res,* **12**, 7663-7675.

Thanos, D. & Maniatis, T. (1992) The high mobility group protein HMG I(Y) is required for NF-κB dependent virus induction of the human IFN-β gene. *Cell,* **71**, 777-789.

Thatcher, T.H. & Gorovsky, M.A. (1994) Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucl Acids Res,* **22**, 174-179.

Thoma, F. & Koller, T.H. (1977) Influence of histone H1 on chromatin structure. *Cell,* **12**, 101-107.

Thoma, F., Koller, T.H. & Klug, A. (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol,* **83**, 403-427.

Thoma, F. & Koller, T. (1981) Unravelled nucleosomes, nucleosome beads and higher order structures of chromatin - influence of non-histone components and histone H1. *J Mol Biol,* **149**, 709-733.

Thoma, F., Losa, R. & Koller, T. (1983) Involvement of the domains of histones H1 and H5 in the structural organization of soluble chromatin. *J Mol Biol,* **167**, 619-640.

Thoma, F., Bergman, L.W. & Simpson, R.T. (1984) Nuclease digestion of circular TRPARS1 chromatin reveals positioned nucleosomes separated nuclease sensitive regions. *J Mol Biol,* **177**, 715-733.

Thoma, F. & Simpson, R.T. (1985) Local protein-DNA interactions may determine nucleosome positions on yeast plasmids. *Nature,* **315**, 250-252.

Thoma, F. (1986) Protein-DNA interactions and nuclease sensitive regions determine nucleosome positions on yeast plasmid chromatin. *J Mol Biol,* **190**, 177-190.

Thoma, F. & Zatchej, M. (1988) Chromatin folding modulates nucleosome positioning in yeast minichromosomes. *Cell,* **55**, 945-953.

Thoma, F. (1991) Structural-changes in nucleosomes during transcription - strip, split or flip. *Trends Genet,* **7**, 175-177.

Thoma, F. (1992) Nucleosome positioning. *Biochim Biophys Acta,* **1130**, 1-19.

Thomas, J.O. & Kornberg, R.D. (1975) An octamer of histones in chromatin and free in solution. *Proc Natl Acad Sci USA,* **72**, 2626-2630.

Thomas, J.O. & Rees, C. (1983) Exchange of histone H1 and histone H5 between chromatin fragments - a preference of H5 for higher order structures. *Eur J Biochem,* **134**, 109-115.

Thomas, J.O. & Wilson, C.M. (1986) Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5. *EMBO J,* **5**, 3531-3537.

Thompson, J.S., Johnson, L.M. & Grunstein, M. (1994) Specific repression of the yeast silent mating locus hmr by an adjacent telomere. *Mol Cell Biol,* **14**, 446-455.

Thorne, A.W., Kmiciek, D., Mitchelson, K., Sautiere, P. & Crane-Robinson, C. (1990) Patterns of histone acetylation. *Eur J Biochem,* **193**, 701-713.

Travers, A.A. & Klug, A. (1987) The bending of DNA in nucleosomes and its wider implications. *Phil Trans Royal Soc London Series B- Biological Sciences,* **317**, 537-561.

Travers, A. (1989) DNA structure - curves with a function. *Nature,* **341**, 184-185.

Travers, A.A. & Muyldermans, S.V. (1996) A DNA-sequence for positioning chromatosomes. *J Mol Biol,* **257**, 486-491.

Tremethick, D.J. & Molloy, P.L. (1988) Effects of high mobility group protein-1 and protein-2 on initiation and elongation of specific transcription by RNA polymerase II invitro. *Nucl Acids Res,* **16**, 11107-11123.

Trifonov, E.N. (1980) Sequence-dependent deformational anisotropy of chromatin DNA. *Nucleic Acids Res* **8**, 4041-4053.

Truss, M., Bartsch, J., Schelbert, A., Hache, R.J.G. & Beato, M. (1995) Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter *in vivo*. *EMBO J.* **14**, 1737-1751.

Tsukiyama, T., Becker, P.B. & Wu, C. (1994) ATP-dependent nucleosome disruption at a heat shock promoter mediated by binding of GAGA transcription factor. *Nature,* **367**, 525-532.

Tsukiyama, T. & Wu, C. (1995) Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell,* **83**, 1011-1020.

Turnell, W.G. & Travers, A.A. (1992) Algorithms for prediction of histone octamer binding sites. *Methods Enzymol,* **212**, 387-399.

Turner, B.M. & O'Neill, L.P. (1995) Histone acetylation in chromatin and chromosomes. *Seminars In Cell Biology,* **6**, 229-236.

Uemura, T., Ohkura, H., Adachi, Y., Morino, K., Shiozaki, K. & Yanagida, M. (1987) DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in s-pombe. *Cell,* **50**, 917-925.

Ulyanov, A.V. & Stormo, G.D. (1995) Multi alphabet consensus algorithm for identification of low specificity protein-DNA interactions. *Nucl Acids Res,* **23**, 1434-1440.

Ulyanov, N.B. & Zhurkin, V.B. (1984) Sequence-dependent anisotropic flexibility of B-DNA - a conformational study. *J Biomol Struc Dynam,* **2**, 361-385.

Ura, K., Hayes, J.J. & Wolffe, A.P. (1995) A positive role for nucleosome mobility in the transcriptional activity of chromatin templates - restriction by linker histones. *EMBO J,* **14**, 3752-3765.

Ura, K. & Wolffe, A.P. (1996) Reconstruction of transcriptionally active and silent chromatin. *Methods Enzymol,* **274**, 257-271.

van Holde, K.E. (1989) Chromatin. Springer-Verlag, New York,

Varga-Weisz, P., van Holde, K. & Zlatanova, J. (1994) Competition between linker histones and HMG1 for binding to four-way junction DNA: implications for transcription. *Biochem Biophys Res commun,* **203**, 1904-1911.

VargaWeisz, P.D., Blank, T.A. & Becker, P.B. (1995) Energy-dependent chromatin accessibility and nucleosome mobility in a cell-free system. *EMBO J,* **14**, 2209-2216.

Vettesse-Dadey, M., Walter, P., Chen, H., Juan, L.J. & Workman, J.L. (1994) Role of the histone amino termini in faciliated binding of a transcription factor, GAL4-AH, to nucleosome cores. *Mol Cell Biol,* **14**, 970-981.

Vettese-Dadey, M., Grant, P.A., Hebbes, T.R., CraneRobinson, C., Allis, C.D. & Workman, J.L. (1996) Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA *in vitro*. *EMBO J,* **15**, 2508-2518.

Voorma, H.O. (1996) Control of translation initiation in prokayotes. In: *Translational control*, edited by Hershey, J.W.B., Mathews, M.B. & Sonenberg, N. Cold Spring Harbor Laboratory Press, New York, p. 759-777.

Wade, P.A., Pruss, D. & Wolffe, A.P. (1997) Histone acetylation: chromatin in action. *Trends Biochem Sci,* **22**, 128-132.

Wagner, L.A., Gesteland, R.F., Dayhuff, T.J. & Weiss, R.B. (1994) An efficient Shine-Dalgarno sequence but not translation is necessary for lacz mRNA stability in *Escherichia coli. J Bacteriol,* **176**, 1683-1688.

Wall, G., Varga-Weisz, P.D., Sandaltzopoulos, R. & Becker, P.B. (1995) Chromatin remodeling by GAGA factor and heat shock factor at the hypersensitive *Drosophila* hsp26 promoter *in vitro. EMBO J,* **14**, 1727-1736.

Walmsley, M.E., Buckle, R.S., Allan, J. & Patient, R.K. (1991) A chicken red cell inhibitor of transcription associated with the terminally differentiated state. *J Cell Biol,* **114**, 9-19.

Walter, P.P., Owen-Hughes, T.A., Cote, J. & Workman, J.L. (1995) Stimulation of transcription factor binding and histone displacement by nucleosome assembly protein-1 and nucleoplasmin requires disruption of the histone octamer. *Mol Cell Biol,* **15**, 6178-6187.

Wang, J., Hogan, M. & Austin, R.H. (1982) DNA motions in the nucleosome core particle. *Proc Natl Acad Sci USA -Biological Sciences,* **79**, 5896-5900.

Wang, S.X. & Hazelrigg, T. (1994) Implications for bcd mRNA localization from spatial distribution of exu protein in *Drosophila* oogenesis. *Nature,* **369**, 400-403.

Ward, W.W., Cody, C.W., Hart, R.C. & Cormier, M.J. (1980) Spectrophotometric identity of the energy transfer chromophores in *Renilla* and *Aequorea* green fluorescent proteins. *Photochem Photobiol,* **31**, 611-615.

Way, M., Pope, B., Gooch, J., Hawkins, M. & Weeds, A.G. (1990) Identification of a region in segment 1 of gelsolin critical for actin binding. *EMBO J,* **9**, 4103-4109.

Wechsler, D.S., Papoulas, O., Dang, C.V. & Kingston, R.E. (1994) Differential binding of c-myc and max to nucleosomal DNA. *Mol Cell Biol,* **14**, 4097-4107.

Weller, J.W. & Hill, W.E. (1992) Probing dynamic changes in rRNA conformation in the 30S subunit of the *Escherichia coli* ribosome. *Biochemistry,* **31**, 2748-2757.

Wellman, S.E., Sittman, D.B. & Chaires, J.B. (1994) Preferential binding of H1e histone to GC-rich DNA. *Biochemistry,* **33**, 384-388.

Widom, J., Finch, J.T. & Thomas, J.O. (1985) Higher order structure of long repeat chromatin. *EMBO J,* **4**, 3189-3194.

Widom, J. & Klug, A. (1985) Structure of the 300 Å chromatin filament: X-ray diffraction from oriented samples. *Cell,* **43**, 207-213.

Williams, S.P., Athey, B.D., Muglia, L.J., Schappe, R.S., Gough, A.H. & Langmore, J.P. (1986) Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys J,* **49**, 233-248.

Williamson, R. (1970) Properties of rapidly labelled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *J. Mol. Biol.*, **51**, 157-168.

Wilson, C.J., Chao, D.M., Imbalzano, A.N., Schnitzler, G.R., Kingston, R.E. & Young, R.A. (1996) RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodelling. *Cell,* **84**, 235-244.

Wilson, I.W., Praszkier, J. & Pittard, A.J. (1994) Molecular analysis of RNA I control of repb translation in incb plasmids. *J Bacteriol,* **176**, 6497-6508.

Wisniewski, J.R. & Schulze, E. (1994) High-affinity interaction of dipteran high mobility group (HMG) proteins-1 with DNA is modulated by C-terminal regions flanking the HMG box domain. *J Biol Chem,* **269**, 10713-10719.

Woese, C.R., Winker, S. & Gutell, R.R. (1990) Architecture of rRNA - constraints on the sequence of tetra loops. *Proc Natl Acad Sci USA,* **87**, 8467-8471.

Wolffe, A.P. (1991) Implications of DNA replication for eukaryotic gene expression. *J Cell Sci,* **99**, 201-206.

Wolffe, A.P., Khochbin, S. & Dimitrov, S. (1997) What do linker histones do in chromatin? *BioEssays,* **19**, 249-255.

Woodcock, C.L.F., Frado, L.L.Y. & Rattner, J.B. (1984) The higher order structure of chromatin - evidence for a helical ribbon arrangement. *J Cell Biol,* **99**, 42-52.

Woodcock, C.L., Grigoryev, S.A., Horowitz, R.A. & Whitaker, N. (1993) A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. *Proc Natl Acad Sci USA,* **90**, 9021-9025.

Woodcock, C.L. (1994) Chromatin fibers observed *in situ* in frozen hydrated sections -native fiber diameter is not correlated with nucleosome repeat length. *J Cell Biol,* **125**, 11-19.

Worcel, A., Strogatz, S. & Riley, D. (1981) Structure of chromatin and the linking number of DNA. *Proc Natl Acad Sci USA-Biological Sciences,* **78**, 1461-1465.

Workman, J.L. & Kingston, R.E. (1992) Nucleosome core displacement *in vitro* via a metastable transcription factor nucleosome complex. *Science,* **258**, 1780-1784.

Wu, C. (1980) The 5′ end of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature,* **286**, 854-860.

Wu, H.M. & Crothers, D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature,* **308**, 509-513.

Wu, T.C. & Lichten, M. (1994) Meiosis induced double-strand break sites determined by yeast chromatin structure. *Science,* **263**, 515-518.

Xu, C.F., Brown, M.A., Chambers, J.A., Griffiths, B., Nicolai, H. & Solomon, E. (1995) Distinct transcription start sites generate 2 forms of brca1 mRNA. *Human Mol Genet,* **4**, 2259-2264.

Yakoto, H., van den Engh, G., Hearst, J.E., Sachs, R.K. & Trask, B.J. (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G_0/G_1 interphase nucleus. *J Cell Biol,* **130**, 1239-1249.

Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H. & Nakatani, Y. (1996) A p300/cbp-associated factor that competes with the *Adenoviral* oncoprotein e1a. *Nature,* **382**, 319-324.

Yao, J., Lowary, P.T. & Widom, J. (1990) Direct detection of linker DNA bending in defined length oligomers of chromatin. *Proc Natl Acad Sci USA,* **87**, 7603-7607.

Yao, J., Lowary, P.T. & Widom, J. (1991) Linker DNA bending by the core histones of chromatin. *Biochemistry,* **30**, 8408-8414.

Yenidunya, A., Davey, C., Clark, D., Felsenfeld, G. & Allan, J. (1994) Nucleosome positioning on chicken and human globin gene promoters *in vitro* novel mapping techniques. *J Mol Biol,* **237**, 401-414.

Yoon, Y.S., Kim, J.W., Kang, K.W., Kim, Y.S., Choi, K.H. & Joe, C.O. (1996) Poly (ADP-ribosyl)ation of histone H1 correlates with internucleosomal DNA fragmentation during apoptosis. *J Biol Chem,* **271**, 9129-9134.

Zentgraf, H. & Franke, W.W. (1984) Differences of supranucleosomal organization in different kinds of chromatin: Cell type specific globular subunits containing different numbers of nucleosomes. *J Cell Biol,* **99**, 272-285.

Zhang, J. & Deutscher, M.P. (1989) Analysis of the upstream region of the *Escherichia coli* rnd gene encoding RNase D - evidence for translational regulation of a putative tRNA processing enzyme. *J Biol Chem,* **264**, 18228-18233.

Zhang, J.R. & Deutscher, M.P. (1992) A uridine-rich sequence required for translation of prokaryotic mRNA. *Proc Natl Acad Sci USA,* **89**, 2605-2609.

Zhurkin, V.B. (1985) Sequence-dependent bending of DNA and phasing of nucleosomes. *J Biomol Struc Dynam,* **2**, 785-804.

Zhurkin, V.B., Ulyanov, N.B., Gorin, A.A. & Jernigan, R.L. (1991) Static and statistical bending of DNA evaluated by Monte-Carlo simulations. *Proc Natl Acad Sci USA,* **88**, 7046-7050.

Zlatanova, J. & Doenecke, D. (1994) Histone H1 degrees - a major player in cell-differentiation. *Faseb J,* **8**, 1260-1268.

Zweidler, A. (1984) Core histone variants of the mouse: primary structure and differential expression. In: *Histone genes*, edited by Stein, G.S., Stein, J.L. & Marzluff, W.F. John Wiley & sons, New York, p. 339-369.